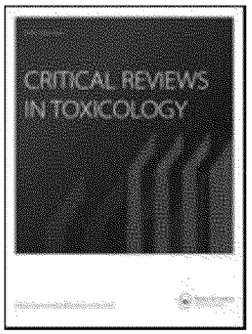

Taylor & Francis
Taylor & Francis Group

Critical Reviews in Toxicology

ISSN: 1040-8444 (Print) 1547-6898 (Online) Journal homepage: http://www.tandfonline.com/loi/itxc20

Quantitative cancer risk assessment for occupational exposures to asphalt fumes during built-up roofing asphalt (BURA) operations

Lorenz R. Rhomberg, David B. Mayfield, Julie E. Goodman, Eric L. Butler, Marc A. Nascarella & Daniel R. Williams

To cite this article: Lorenz R. Rhomberg, David B. Mayfield, Julie E. Goodman, Eric L. Butler, Marc A. Nascarella & Daniel R. Williams (2015) Quantitative cancer risk assessment for occupational exposures to asphalt fumes during built-up roofing asphalt (BURA) operations, Critical Reviews in Toxicology, 45:10, 873-918, DOI: 10.3109/10408444.2015.1094450

To link to this article: http://dx.doi.org/10.3109/10408444.2015.1094450

CRITICAL REVIEWS IN TOXICOLOGY, 2015
VOL. 45, NO. 10, 873–918
http://dx.doi.org/10.3109/10408444.2015.1094450

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

Quantitative cancer risk assessment for occupational exposures to asphalt fumes during built-up roofing asphalt (BURA) operations

Lorenz R. Rhomberg[1], David B. Mayfield[2], Julie E. Goodman[1], Eric L. Butler[1], Marc A. Nascarella[1] & Daniel R. Williams[1]

[1]Gradient, Cambridge, MA, USA and [2]Gradient, Seattle, WA, USA

ABSTRACT

The International Agency for Research on Cancer qualitatively characterized occupational exposure to oxidized bitumen emissions during roofing as probably carcinogenic to humans (Group 2A). We examine chemistry, exposure, epidemiology and animal toxicity data to explore quantitative risks for roofing workers applying built-up roofing asphalt (BURA). Epidemiology studies do not consistently report elevated risks, and generally do not have sufficient exposure information or adequately control for confounders, precluding their use for dose–response analysis. Dermal carcinogenicity bioassays using mice report increased tumor incidence with single high doses. In order to quantify potential cancer risks, we develop time-to-tumor model methods [consistent with US Environmental Protection Agency (EPA) dose–response analysis and mixtures guidelines] using the dose–time–response shape of concurrent exposures to benzo[*a*]pyrene (B[*a*]P) as concurrent controls (which had several exposure levels) to infer presumed parallel dose–time–response curves for BURA-fume condensate. We compare EPA relative potency factor approaches, based on observed relative potency of BURA to B[*a*]P in similar experiments, and direct observation of the inferred BURA dose–time–response (scaled to humans) as means for characterizing a dermal unit risk factor. We apply similar approaches to limited data on asphalt-fume inhalation and respiratory cancers in rats. We also develop a method for adjusting potency estimates for asphalts that vary in composition using measured fluorescence. Overall, the various methods indicate that cancer risks to roofers from both dermal and inhalation exposure to BURA are within a range typically deemed acceptable within regulatory frameworks. The approaches developed may be useful in assessing carcinogenic potency of other complex mixtures of polycyclic aromatic compounds.

ARTICLE HISTORY
Received 30 October 2014
Revised 2 September 2015
Accepted 7 September 2015
Published online 27 October 2015

KEYWORDS
Asphalt, benzo[a]pyrene, carcinogenicity, epidemiology, polycyclic aromatic hydrocarbon, relative potency factor, risk assessment, time to tumor

Table of contents

CONTACT Lorenz R. Rhomberg, lrhomberg@gradientcorp.com Gradient, 20 University Road, Cambridge, MA 02138, USA. Tel: +1 6173955000.
Supplemental data for this article can be accessed at http://dx.doi.org/10.3109/10408444.2015.1094450.

Introduction

In 2013, the International Agency for Research on Cancer (IARC) released a monograph describing the carcinogenic potential of bitumen and bitumen emissions (IARC 2013a). In North America, bitumen is more commonly called asphalt; thus, in this paper, we use the term asphalt – or, when referring specifically to asphalt used in roofing applications, built-up roofing asphalt (BURA). In this paper, we focus on hot-applied BURA, which comprises one among a number of roofing products (including soft- or cold-applied asphalt materials) (ARMA et al. 2011). Asphalts are materials derived from crude oil that are used for road paving and roofing materials (IARC 2013a). When heated, asphalts release mixtures of aerosols and gases (i.e. asphalt fumes). Petroleum-derived products, such as asphalts, contain potential and probable carcinogenic polycyclic aromatic hydrocarbons (PAHs) (IARC 2013a). Based on the presence of carcinogenic PAHs, IARC evaluated several types of asphalt products. One subgroup included oxidized asphalt, which is produced by blowing air through hot asphalt. IARC classified occupational exposure to oxidized asphalt and its emissions during roofing to be probably carcinogenic to humans (Group 2A) based on limited evidence in humans, limited evidence in animals for oxidized asphalt and sufficient evidence in animals for fume condensates of oxidized asphalt. Hot application of BURA to make low slope roofs is included in this classification.

Although IARC (2013a) considers cancer hazard qualitatively, the monograph does not provide a quantitative characterization of cancer risks associated with exposure to BURA or BURA emissions. In this respect, IARC's cancer assessment provides limited information to inform a dose–response assessment to evaluate the potential risks to asphalt workers. In this paper, we focus on exploring alternative methods to define a basis for quantifying cancer risk in roofing workers exposed to BURA and BURA emissions. We focus particularly on ways to make quantitative risk projections for roofing workers using the data that IARC relied upon to arrive at its qualitative classification of bitumen's potential carcinogenicity. Although there are larger, complex and varied questions about carcinogenicity of PAHs in general, and about how observations of the potencies of some mixtures might inform estimates of the potencies of different PAH mixtures, we have not attempted to review and synthesize interpretations of the entire field. Rather, we use these larger questions as a basis for gauging the quantitative risk implications for actual roofing worker exposures in studies of roofing asphalt fume condensates on which IARC based its qualitative carcinogenicity classification.

Evaluation of PAH mixtures is complicated by the large suite of compounds known to be present, the relative lack of hazard information and the uncertainties with extrapolating data from individual components to whole mixtures (Jarvis et al. 2014; NTP 2012). Owing to these knowledge gaps, current PAH risk assessment practices are often limited to a subset of mixture components [such as the US Environmental Protection Agency's (EPA's) 16 priority PAHs]. This restriction often entails the use of relative potency factors (RPFs) to extrapolate across individual PAHs and necessitates broad assumptions about mechanisms of action and dose–response relationships. Various sets of RPFs exist (although individual RPFs are not available for all PAHs); they often differ in magnitude of presumed potency, and some sets are under revision (see review by Jarvis et al. 2014). Further PAH testing strategies have been proposed to enhance the understanding of PAHs by examining toxicity (through *in vivo* and *in vitro* studies), toxicokinetics (to inform route-to-route extrapolation), and exposure estimation (to reflect actual human exposures) (Jarvis et al. 2014; NTP 2012). In this paper, we examine various techniques for adjusting the assessment of PAH mixtures' potency based on assessment of changes in mixture composition (i.e. specifically BURA from different sources or used at different temperatures). To enable this risk assessment, we have relied upon what data are available for the subject PAH mixtures. We recognize that this assessment is limited by the uncertainties in current information, RPFs, and risk assessment practices. We nonetheless have explored several different and complementary analyses, focusing when possible on directly assessing whole-mixture data but also considering these observed potencies relative to those of concurrent benzo[*a*]pyrene (B[*a*]P) positive controls and estimated contributions of individual mixture components.

One specific challenge in assessing the toxicity of complex mixtures is that the contributions of potentially toxic components of a mixture could vary from case to case, and some components could potentially affect others' toxicity in non-additive ways, either enhancing or inhibiting effects. The EPA's mixtures guidelines (EPA 2000) provide a structured approach to address this challenge. Our paper focuses on assessing compositional variability and dose–response techniques to address mixture similarities, whereas other investigators have

employed alternative dose–response techniques (Gennings et al. 2005) or pharmacokinetic modeling approaches (Haddad et al. 2000; Tan et al. 2011).

EPA guidance for assessment of chemical mixtures (EPA 2000) suggests that, when possible, exposure to a multi-chemical mixture of reasonably consistent composition should be evaluated using direct measures of the subject mixture's toxicity rather than that of its individual components. The key questions when following this approach are whether the mixture is indeed consistent in composition, whether individual samples are representative of the mixture for which toxicity observations are available, and whether the toxicity data are sufficient to characterize its effects. If data on the subject mixture are deemed insufficient, or, if its composition varies sufficiently from sample to sample, then toxicity observations on a suitably similar mixture can be used. Whether a mixture is deemed suitably similar depends partly on how greatly concentrations of component chemicals vary and to what degree their toxicity is described. Attempting to determine the toxicity consequences of this variation in mixture composition is important, because doing so may permit adjustment for relative toxic potency of similar mixtures that share an underlying mode of toxic action. Finally, if one cannot base an assessment on a sufficiently similar mixture, EPA (2000) recommends basing the assessment on knowledge of individual components' toxicity, assuming that the component-specific risks combine to produce the effect of the whole. The challenge for this approach is to characterize the individual components sufficiently and to assess whether the components that have differing potencies comprise the whole basis for the mixture's toxicity (i.e. whether other components without individual potency information might also make important contributions to the overall effect). It has been common practice to use the RPF approach for risk assessments of PAH mixtures – that is, combining risks from component PAHs in the mixture, with each component's contribution measured by a potency relative to that of B[*a*]P (EPA 1993, 2010). In the case of BURA, there are ample data to describe chemical compositions (e.g. Cavallari et al. 2012a, b; Kriech et al. 2007; Machado et al. 1993). Data are limited, although useful, on occupational exposures (e.g. Kriech et al. 2004a; McClean et al. 2007), dose–response relationships from animal studies (e.g. Clark et al. 2011; Freeman et al. 2011; Fuhst et al. 2007; Sivak et al. 1997), and epidemiology associations (e.g. Boffetta et al. 2003a, b; Fayerweather 2007; Partanen & Boffetta 1994). The analysis presented here explores the utility of existing data for estimating the cancer risk from exposure to BURA in current North American roofing applications.

Our aim is to assess the carcinogenic potency of BURA (and emissions) as experienced by roofing workers (in a manner consistent with current EPA dose–response analysis and mixtures assessment guidelines). In the analysis that follows, we explore several means of deriving such potency estimates, guided by the considerations described above for EPA's mixtures assessment. The purpose of considering alternative approaches is to gauge their respective strengths and shortcomings, to articulate their inherent assumptions, and to assess the consistency of outcomes among methods (see Figure 1 for components of this analysis). The following sections present the existing data and subsequent analyses separately for inhalation and dermal exposure to emissions from heated asphalt mixtures (i.e. BURA and paving asphalts).

Our study exists in the context of a much larger set of questions, discussions and scientific debates about PAH carcinogenicity and its many aspects: differences in mode of action and potency among PAH constituents, interactions between PAHs, exposure route-specific absorption, metabolism and effects, differences between rodent and human sensitivity to PAHs, and other similarly fundamental questions. We acknowledge that contending approaches to the factors described above could affect our analyses, and that we cannot hope to solve all of the problems in this larger context. We have focused on the studies that led to roofing asphalt fumes being deemed a human carcinogen and explored the implications of these studies, unaddressed till now, for quantitative assessment of possible risks to roofing workers. Although we are not endorsing the IARC (2013a) analysis, and while we acknowledge its critics, we take its qualitative findings on asphalt fumes' cancer risk potential as a starting point. Similarly, the EPA's draft assessment of B[*a*]P carcinogenicity (EPA 2014a, b), which quantitatively evaluates its cancer potency by different routes of exposure, is an important part of the context of our analyses. This assessment is still in draft form and may change, and public commenters have leveled a number of potent criticisms of its current state. Some of our approaches use observations of BURA-fume condensate and B[*a*]P cancer potencies in the same experimental system as a means for defining BURA potency relative to that of B[*a*]P; this relative potency could be used with an externally supplied potency for B[*a*]P to project human cancer risk levels. We use EPA's (2014a, b) current draft B[*a*]P potencies to make these calculations specific, but our own analysis concerns only the estimated relative potency of BURA and B[*a*]P. These ratios could be applied to quantify different baseline estimates of B[*a*]P potency, either revised ones from EPA or estimates from other sources.

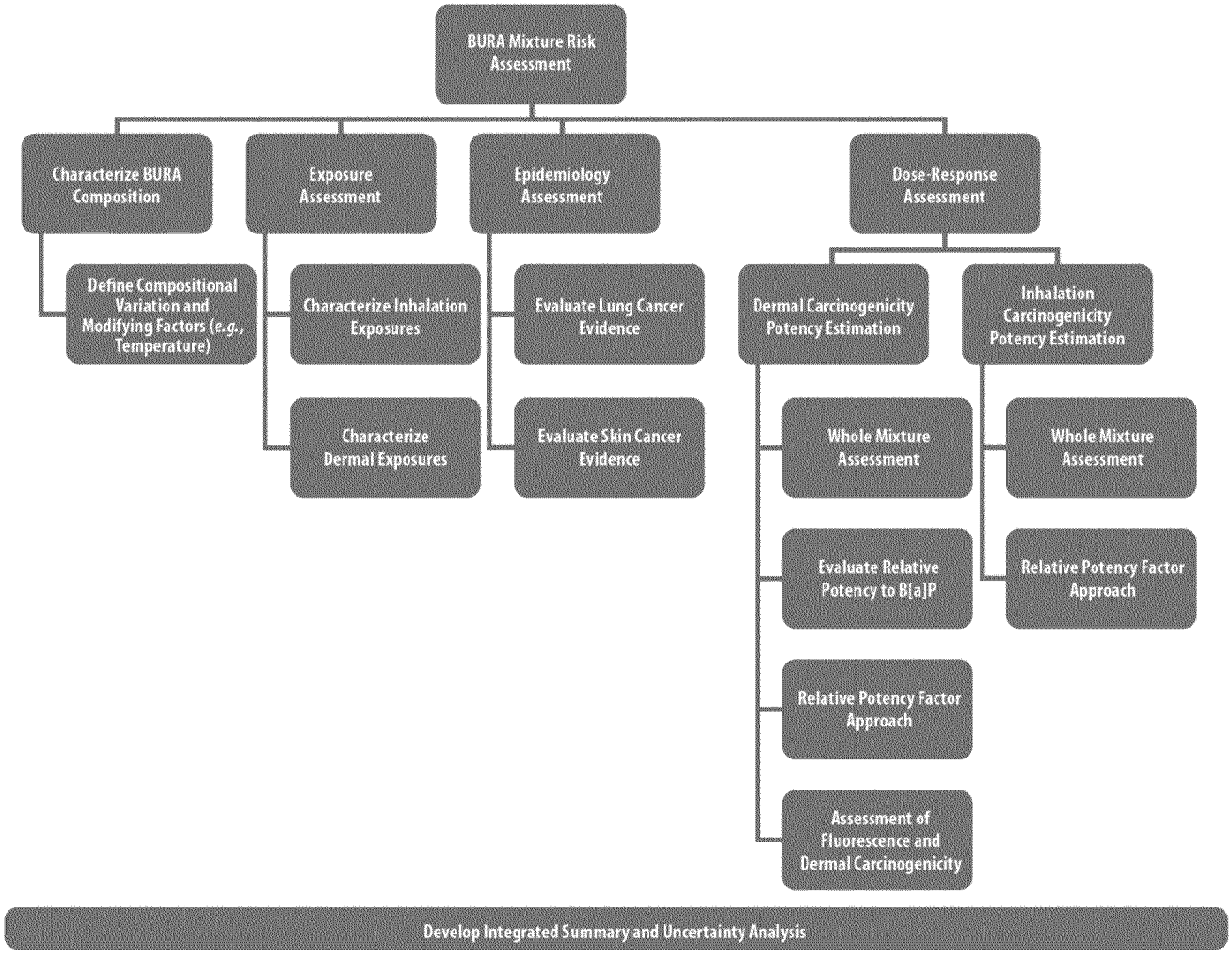

Figure 1. Flow chart for the BURA risk assessment analysis.

Asphalt chemistry and composition

We reviewed asphalt chemistry to evaluate whether BURA fume composition is consistent (i.e. to determine whether the BURA fume condensate is a consistent mixture). Per EPA's (2000) mixture guidelines, chemical compositions should be examined in relation to the variability associated with alternative analytical techniques, alternative sources (e.g. emission sources or product sources) and transformations between environmental compartments (e.g. fumes generated when the asphalt mix is heated). Characterization of the mixture composition under varying environmental conditions will help in interpreting observations of exposure and toxicity.

Asphalt is derived from residuum of vacuum-reduced distillation of crude oils. Crude oil is a complex mixture of organic material, mostly carbon and hydrogen, with smaller amounts of sulfur, oxygen, nitrogen and trace metals, such as vanadium, nickel, iron and copper (Speight 1999). The chemical composition proportions depend on the source of the crude (Speight 1999). Further processing by air blowing alters its physical characteristics (e.g. oftening points, viscosity) to make materials resilient to temperature changes and weather conditions. BURA materials are typically organized by industry specifications (Type I-IV) for appropriate roof pitch angles in North America (API 2009), but not in Europe (ARMA et al. 2011). This evaluation focuses on studies using BURA Type III, the most common BURA product used in North America (ARMA et al. 2011). Two categories (oxidized and straight run) comprise 99% of asphalt materials used in the USA and Europe (API 2009). Oxidized asphalt (CAS No. 64742-93-4), also called air-blown asphalt, is typically used in BURA application as well as to make asphalt shingles and for various industrial applications (API 2009). Oxidized asphalt contains mostly aliphatic hydrocarbons, PAHs and heterocyclic compounds, in which one or more carbons in the ring compound have been replaced with an atom of nitrogen, oxygen or sulfur (IOMC 2004). Oxidized asphalts are prepared from asphalt feedstock and

Table 1. Concentration of PAHs in heated (316 °C) BURA fumes and coal tar pitch fumes.

PAH[a]	Roofing asphalt (Fume [ppm])[b]	Coal tar pitch (Fume [ppm])
Acenaphthene	12	960
Acenaphthylene	ND	240
Anthracene	0.6	500
Benz[*a*]anthracene	2.6	390
Benzo[*a*]pyrene	1.6	480
Benzo[*b*]fluoranthene	1.2	400
Benzo[*g*,*h*,*i*]perylene	5.4	285
Benzo[*k*]fluoranthene	0.4	110
Chrysene	19.4	1620
Dibenz[*a*,*h*]anthracene	ND	60
Fluoranthene	0.6	360
Fluorene	28.4	1560
Indeno[1,2,3-*c*,*d*]pyrene	ND	90
Naphthalene	39.2	2460
Phenanthrene	8.4	1090
Pyrene	8.5	1420
Sum of PAHs	128.3	12 025

BURA, built-up roofing asphalt; ND, non-detected; PAH, polycyclic aromatic hydrocarbon; ppm, parts per million.
[a]Results adapted from Machado et al. (1993).
[b]BURA Type III (air-blown).

Table 2. Variability of PAHs in BURA and paving fumes.

PAH[a,b,c,d]	Ring structure	Concentration (ppm)			
		Min	Max	Mean	SD
Acenaphthene	2-ring	<0.08	240	34.4	56.9
Acenaphthylene	2-ring	<0.08	0.3	0.08	0.15
Fluorene	2-ring	0.2	320	65.2	92.7
Naphthalene	2-ring	<0.08	370	67.1	103
Anthracene	3-ring	0.4	76	18.3	23
Fluoranthene	3-ring	<0.08	89	16.4	18.3
Phenanthrene	3-ring	2.9	799	201	210
Benz[*a*]anthracene	4-ring	<0.08	36	12.3	11.2
Benzo[*b*]fluoranthene	4-ring	<0.08	54	8.58	12
Benzo[*k*]fluoranthene	4-ring	<0.08	5.4	1.89	1.32
Chrysene	4-ring	0.48	264	56.8	57.8
Pyrene	4-ring	1.7	180	43.3	45.8
Benzo[*a*]pyrene	5-ring	<0.08	18	4.67	4.1
Benzo[*e*]pyrene	5-ring	<0.08	17	6.28	6.13
Dibenz[*a*,*h*]anthracene	5-ring	<0.08	0.9	0.08	0.25
Indeno[1,2,3-*c*,*d*]pyrene	5-ring	<0.08	8.9	2.75	3.07
Benzo[*g*,*h*,*i*]perylene	6-ring	<0.08	15	4.54	4.33

BURA, built-up roofing asphalt; Max, maximum; Min, minimum; PAH, polycyclic aromatic hydrocarbon; SD, standard deviation.
[a]Data summarized from: Machado et al. (1993), Brandt et al. (1985), Brandt and De Groot (1996) and Kriech et al. (2007).
[b]Asphalt fume generated in the laboratory, temperature varies by study and sample (temperature range: 147–316 °C).
[c]EPA priority list of 16 PAHs and benzo[*e*]pyrene.
[d]Analytical methods: EPA Method 8310, HPLC-UV Fluorescence; EPA SW-846 Method 8270C, GC/MS.

subjected to air-blowing to meet industry specifications (ARMA et al. 2011). Straight run asphalt (CAS No. 8052-42-4) is used in road paving or roofing applications.

In comparison, coal tar or coal tar pitch (produced from bituminous coal destructive distillation) has far greater concentrations of PAHs than oxidized or straight run asphalt (API 2009). Use of coal tar materials in roofing applications has declined steadily in recent decades, comprising less than 1% of the low-slope roofing market in the USA (ARMA et al. 2011). Comparatively, PAH concentrations in coal tar pitch fume exceed those in BURA fume, with B[*a*]P concentrations approximately 300 times greater in the coal tar pitch fume (Table 1).

The oxidation process affects PAH composition (Bolliet et al. 2013; Trumbore et al. 2011). The total 4- to 6-ring PAH content of oxidized asphalt was reported to decrease (27–60%) when compared to straight run asphalt (Trumbore et al. 2011). Bolliet et al. (2013) examined 24 asphalt samples from different European refineries for PAH composition. Bolliet et al. compared various feedstock and matching blown asphalt samples (air-rectified and oxidized) processed with and without flux oil. Oxidized asphalt (without flux oil) had the lowest PAH concentrations (sum of 16 PAHs = 13.6 mg/kg), while the samples with flux oil had the highest PAH concentrations (sum of 16 PAHs = 43.1 mg/kg). Thus, asphalt is inherently low in PAHs, and processing by air blowing further reduces the PAH concentrations (Bolliet et al. 2013; Trumbore et al. 2011).

Typical BURA application at a worksite involves heating the asphalt on site. The roofing worker places the asphalt in a heating kettle and heats the material to prepare it for application. During heating, the hot asphalt emits fume as vapors and aerosols containing PAHs. Heating asphalt at low temperatures emits substantially less asphalt fume (by mass) and lower concentrations of potentially carcinogenic PAHs, especially 4- to 6-ring compounds (Brandt & De Groot 1996; Machado et al. 1993). As a result, kettle temperature is an important factor when determining potential occupational exposure during roofing application.

A number of publications have assessed the variability of PAH concentrations in BURA and paving asphalt samples from fumes generated in the laboratory (Brandt & De Groot 1996; Brandt et al. 1985; Kriech et al. 2007; Machado et al. 1993). The concentrations of 2- to 3-ring PAHs were greater in magnitude and were more variable than those of 4- to 6-ring PAHs in BURA fume samples (Table 2). For example, BURA and paving asphalt fume condensates had a concentration range of 2.9–799 mg/kg for phenanthrene (a 3-ring PAH), while sample concentrations for B[*a*]P (a 5-ring PAH) were <0.08–18 mg/kg. Overall, the results in Table 2 illustrate the modest variability for concentration of PAHs in BURA and paving asphalt fumes.

The variation in PAH composition [geometric mean (GM) PAH concentrations for 20 paving samples and 5 BURA samples] under varying temperatures is summarized in Table 3. Concentrations of PAHs increased one- to sixfold (depending on ring structure) with increasing temperatures. B[*a*]P concentrations in BURA and paving

samples (heated at 230–315 °C) ranged from 0.42 to 2.5 μg/m^3 and 0.49 to 1.51 μg/m^3, respectively. The data in Table 3 (presented in Figure 2) indicate that asphalt products heated at lower temperatures (i.e. 230 °C) tend to preferentially emit "lighter" compounds (i.e. low molecular weight, 2- to 3-ring PAHs) and fewer "heavy" compounds (i.e. high molecular weight, 4- to 6-ring compounds). In contrast, fumes generated at higher temperatures (i.e. 315 °C) emit slightly increased amounts of 4- to 6-ring compounds (Figure 2). Asphalt fume PAH concentrations increase overall as temperature increases. Specifically, higher concentrations of low molecular weight PAHs (2–3 rings) than of 4- to 6-ring PAHs are generated at lower temperatures. As the temperature increases from 230 to 315 °C, the amount (mass) of fume emitted increases (Figure 3), but the concentration of the PAHs being emitted increases only slightly – approximately 2–3 times, except for chrysene (Table 3).

Table 3. PAH composition of BURA and paving fume samples under varying temperature conditions.

PAH	Ring structure	GM concentration (μg/m^3)[a] BURA samples ($n = 5$) (230 °C)	(315 °C)	Paving samples ($n = 20$) (230 °C)	(315 °C)
Acenaphthene	2-ring	180	402	108	257
Fluorene	2-ring	88.7	158	27.6	50.3
2-Methyl naphthalene	2-ring	167	301	111	198
Naphthalene	2-ring	142	270	81	184
Anthracene	3-ring	141	281	17.2	58.1
Fluoranthene	3-ring	47.3	115	20.6	48.4
Phenanthrene	3-ring	95.8	115	101	186
Benz[*a*]anthracene	4-ring	10.5	33.1	20	40.1
Benzo[*b*]fluoranthene	4-ring	5.56	12.1	4.71	15.9
Benzo[*k*]fluoranthene	4-ring	2.82	8.72	1.53	7.34
Chrysene	4-ring	2.51	15.3	1.46	3.97
Pyrene	4-ring	29.9	72.6	8.86	20.7
Benzo[*a*]pyrene	5-ring	0.42	2.5	0.49	1.51
Benzo[*e*]pyrene	5-ring	1.33	5.62	0.19	0.88
Dibenz[*a,h*]anthracene	5-ring	1.8	4.82	<0.1	0.78
Indeno[1,2,3-*c,d*]pyrene	5-ring	1.72	5.44	0.14	0.59
Benzo[*g,h,i*]perylene	6-ring	1.26	5.24	<0.1	0.8
Sum of 7, 2–3-ring PAHs		861	1642	466	982
Sum of 10, 4–6-ring PAHs		57.8	165	37.4	92.6
Sum of 17, 2–6-ring PAHs		919	1807	503	1075

BURA, built-up roofing asphalt; *n*, sample size; PAH, polycyclic aromatic hydrocarbon.
[a]Results adapted from Cavallari et al. (2012b).

Based on the available information about PAH composition, there is some variation attributable to different source materials; however, only a modest degree of variation in composition exists among sources (i.e. as gauged by PAH concentrations consistently measured across studies). The fume generation temperature is critical, with higher temperatures mildly enriching vapors in the large multi-ringed components that are the chief source of potential carcinogenic effects. Despite this variation, concentrations among components within this group of large multi-ringed PAHs are reasonably consistent, suggesting that increased temperature elevates them as a whole compared to those of smaller, lighter components, but only moderately affects the ratios among the heavier components. Thus, the representativeness of fume-generation processes under field conditions is useful for interpreting the toxicity testing results, in that heated asphalt mixture compositions are relatively consistent, and data from exposure and toxicity studies are generally applicable for characterizing BURA, paving asphalts, and similar asphalt materials. Further, the composition of BURA and paving asphalts are distinctly different from those of other PAH mixtures (i.e. coal tar). Therefore, the analysis provided in the following sections should be reviewed carefully before applying it to other asphalt products, and, accordingly, conclusions for other PAH mixtures should not be used to characterize BURA. Our primary aim is to characterize risks from exposure to BURA Type III emissions; these results can be applied more broadly

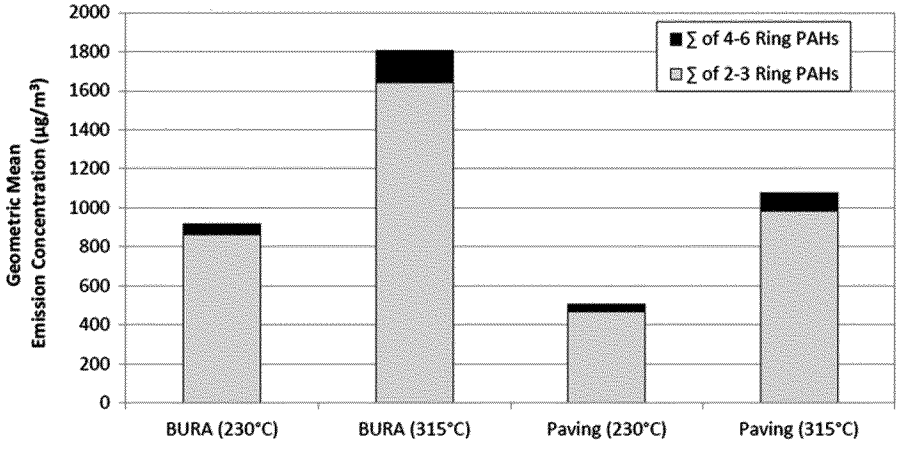

Figure 2. BURA and paving PAHs emission concentration (μg/m^3). *Source*: Adapted from Cavallari et al. (2012a).

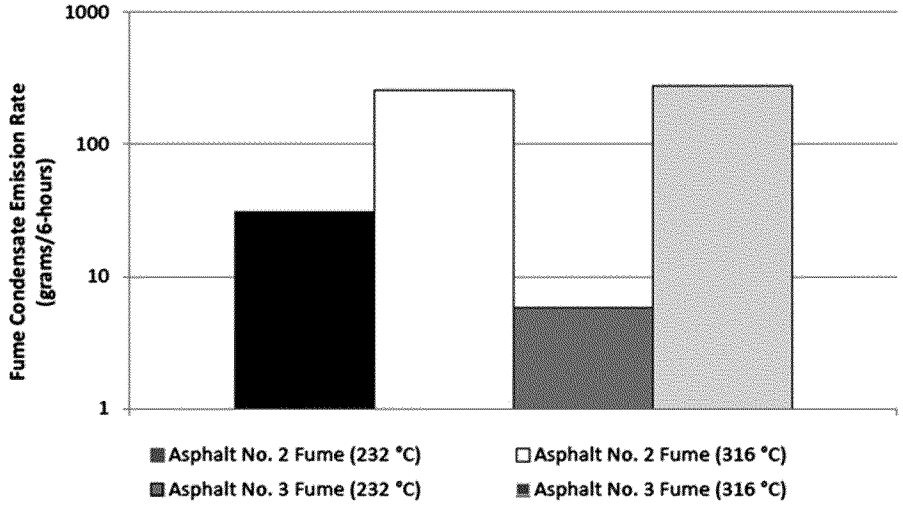

Figure 3. Effect of temperature on asphalt fume condensates emissions. *Source*: Adapted from Machado et al. (1993).

(to similar asphalt materials) as long as care is taken to consider factors that affect compositional variation (e.g. heating temperature), as discussed in the following sections.

Occupational exposure

The exposure assessment in any risk characterization defines the extent to which a population is exposed to the material (or chemical) of concern. When examining mixtures, it is important to evaluate exposure pathways, mixture emissions data and mixture compositions in the environments that match the expected exposure pathways (EPA 2000). In this case, the population of interest is North American roofing workers exposed to BURA fumes and fume condensates. The estimated workforce in the USA (reported in 2008) includes approximately 174 000 workers engaged in roofing construction and removal; roughly 99 000 of these roofers work in the low-slope sector, where hot- and soft-applied asphalt roofing systems are installed (ARMA et al. 2011). Estimates for the European workforce have not been adequately quantified (IARC 2013a), but they are expected to be around the same order of magnitude as the USA (ARMA et al. 2011). In the USA, a small percentage of total roofing hours includes installation of hot-applied asphalt systems (7%) or soft-applied systems (2%) (ARMA et al. 2011). Specialization of workers (by job type) has decreased, and current information is insufficient to quantify what proportion of the overall workforce is dedicated to installation of BURA. In addition, demographic changes have significantly reduced the tenure (formerly 20–30 years) of roofing workers (ARMA et al. 2011). Therefore, the size of the current workforce consistently exposed to BURA (over typical workday or typical work tenure) and their duration of exposure have been declining (ARMA et al. 2011). Occupational studies of workers exposed to oxidized roofing asphalt have been reviewed previously (ARMA et al. 2011; IARC 2013a; NIOSH 2000), and those observations are summarized in the "Occupational inhalation exposure" and "Occupational dermal exposure" sections.

Occupational inhalation exposure

Two primary sampling methods are used to characterize air exposures during roofing asphalt application: one analyzes total particulates, and the other analyzes the benzene-soluble fraction of total particulates (NIOSH 2000). Total particulate matter (TPM) is a measure of all airborne particulates (in the aerosol fraction) that can be collected on a tared (weighed) sample filter, and this measure should be compared to the NIOSH occupational short-term (15 min) exposure limit of 5 mg/m^3 for asphalt fumes (IARC 2013a; NIOSH 2000). The benzene-soluble particulate fraction (BSF) is the portion of total particulates that is soluble in benzene (organic compounds are generally soluble in benzene). The American Conference of Governmental Industrial Hygienists (ACGIH) developed an 8-h time-weighted average (TWA) occupational exposure limit of 0.5 mg/m^3 based on the benzene-soluble inhalable particulate, an approach that is comparable to the BSF method (ACGIH 2001; IARC 2013a; NIOSH 2000). Some investigators have also applied total organic matter (TOM; total organic fraction of both the aerosol and vapor fractions) as a measure of exposure (IARC 2013a; Kriech et al. 2004a, 2010). Asphalt exposure characterization studies in Europe apply an alternative method (IFA Method 6305) that provides higher estimates of TOM than US

Table 4. Summary of field-based personal air sampling for roofing workers.

Exposure measure	Number of exposure studies	Units	All roofing workers[a]					Kettlemen				
			N[b]	Min	Max	GM	AM	N[b]	Min	Max	GM	AM
Total particulate matter	14[c]	mg/m^3 TWA	299	0.04	13	0.9	1.4	53	0.2	13	1.6	2.3
Benzene soluble fraction	18[c]	mg/m^3 TWA	338	0.01	9.6	0.4	0.8	56	0.03	9.6	1.0	1.6
Total organic matter	4[d]	mg/m^3 TWA	115	0.2	8.0	0.7	1.5	20	0.2	11.8	1.6	2.4

AM, arithmetic mean; GM, geometric mean; Max, maximum; Min, minimum; *N*, sample size; TWA, time-weighted average.
[a]All worker types are summarized in this category including kettlemen.
[b]Summary statistics weighted by number of subjects within each exposure study.
[c]Summarized from the following sources: ARMA et al. (2011), Brand et al. (1985), Hicks (1995), Gamble et al. (1999), NIOSH (2000), Toraason et al. (2001) and Kriech et al. (2004a,b, 2007, 2010).
[d]Summarized from the following sources: Gamble et al. (1999) and Kriech et al. (2004a,b, 2010).

methods do (Kriech et al. 2010, as cited by IARC 2013a). Individual PAH measurements in air have also been reported (Brandt et al. 1985; Hicks 1995; Kriech et al. 2007; Wolff et al. 1989); however, the sets of PAH compounds described have been inconsistent.

Reviews of occupational exposure to asphalt fumes during roofing applications (1970–2010) have reported GM concentrations of 0.9 mg/m^3 for TPM (range = 0.04–13 mg/m^3), 0.4 mg/m^3 for BSF (range = 0.01–9.6 mg/m^3) and 0.7 mg/m^3 for TOM (range = 0.2–8.0 mg/m^3) (Table 4). Worker exposure to asphalt emissions has decreased over time (except among kettle operators) with improved safety practices, and recent TPM, BSF and TOM concentrations are generally below occupational standards. Kettlemen consistently experience higher exposures to asphalt emissions, and measured BSF concentrations are sometimes above the threshold limit value (TLV) of 0.5 mg/m^3 (Table 4). The variation in job tasks and descriptions of these tasks in the primary literature complicates interpretation of occupational data (ARMA et al. 2011; NIOSH 2000). Nevertheless, kettle operators frequently encounter higher concentrations of TPM and BSF than other workers, likely due to increased fume generation under elevated temperatures and their proximity to heated asphalt (Kriech et al. 2004a; NIOSH 2000).

Comparing measurements (e.g. TPM, BSF and TOM) are also complicated because different fraction are measured (e.g. particulates vs. fumes), and the potential for confounding exposures (e.g. particles from diesel exhaust) is considerable. Kriech et al. (2004a, b, 2010) conducted analyses of confounders for varying measurement techniques and documented poor correlations between metrics, attributed to confounders. However, stronger correlations were observed among the different measurement techniques when test conditions removed potential confounders (e.g. roof tear-off particulates, diesel exhaust or other construction materials). Kriech et al. (2010) examined the relationships between sampling methods and found significant correlations, such that predictive factors could be used to relate differing methods.

Few occupational studies have characterized PAH concentrations during BURA application (Hatjian et al. 1995, 1997; Hicks 1995; Kriech et al. 2007; Toraason et al. 2001; Wolff et al. 1989). Hatjian et al. (1995) reported total PAH TWA exposure concentrations (eight PAHs quantitated) for 6 pavers and 10 roofers over 3 days of work ranging from 0.012 to 1.47 μg/m^3. A follow-up study (Hatjian et al. 1997) reported a 3-day GM TWA PAH exposure concentration of 0.442 μg/m^3 for 9 roofing workers using oxidized roofing asphalt (nine PAHs quantitated). Wolff et al. (1989) reported mean TWA PAH exposure concentrations (6–7 samples per day) over 2 days of asphalt roofing ranging from 5.8 to 22.3 μg/m^3. Higher PAH concentrations were noted during roof tear-off activities (mean = 13.4–23.0 μg/m^3) due to the presence of coal tar (Wolff et al. 1989). Hicks (1995) reported a mean total TWA PAH exposure concentration of 19 μg/m^3 (based on 11 samples) for roofing contractors. Finally, Toraason et al. (2001) monitored 26 roofing workers (19 with coal tar exposure and 7 with asphalt-only exposure) during roof removal and asphalt application, and the range of total PAHs was reported to be 16.4–198.2 μg/m^3 (the lower end of the range reflects asphalt-only exposures). In these studies, 2- to 3-ring PAHs comprised most of the total PAH concentrations (85%), and 4- to 6-ring PAHs made up a smaller proportion (ARMA et al. 2011).

Kriech et al. (2007) performed a detailed study of paving and roofing asphalt fumes under conditions representative of actual worker exposures. In the first phase of the study, both TPM and BSF concentrations were measured during application of asphalts from four paving or roofing asphalt sources (A, B, C and D). In the second phase of the study, laboratory-generated fumes were compared to field collected fumes to evaluate PAH variation resulting from different fume generation methods (lab vs. field). To obtain an asphalt fume representative of workplace fumes, the roofing tank

Table 5. Individual PAHs detected in large volume laboratory-generated fumes.

PAH	Ring structure	Paving samples (mg/kg)[a]				Roofing samples (mg/kg)[a]				
		TP-A	TP-B	TP-C	TP-D	TR-A	LR-A	TR-B	TR-C	TR-D
Acenaphthene	2-ring	73	80	120	240	16	18	86	19	23
Fluorene	2-ring	150	210	220	320	110	82	180	69	94
Naphthalene	2-ring	190	150	370	250	110	39	210	77	140
Anthracene	3-ring	40	67	74	76	23	33	62	16	11
Fluoranthene	3-ring	7.9	17	23	10	6.8	22	89	11	6.9
Phenanthrene	3-ring	300	450	500	500	130	250	370	70	51
Benz[*a*]anthracene	4-ring	1.1	1.8	4	<0.08	12	20	26	12	9
Benzo[*b*]fluoranthene	4-ring	<0.08	<0.08	0.94	<0.08	2.8	5.5	4.9	2.7	2.1
Benzo[*k*]fluoranthene	4-ring	<0.08	<0.08	1.1	<0.08	1.9	2.9	3.5	1.5	1.7
Chrysene	4-ring	1.8	2	4	0.48	14	20	16	12	11
Pyrene	4-ring	21	17	11	33	25	150	180	14	10
Triphenylene	4-ring	0.44	0.51	1.1	0.16	3.9	5.8	8.6	3.9	3
Benzo[*a*]pyrene	5-ring	<0.08	<0.08	0.55	<0.08	3.1	5.8	7	2.6	1.2
Benzo[*e*]pyrene	5-ring	<0.08	<0.08	0.47	<0.08	8.1	17	12	6.9	5.7
Indeno[1,2,3-*c*,*d*]pyrene	6-ring	<0.08	<0.08	<0.08	<0.08	0.87	1.2	1.1	0.74	0.63
Benzo[*g*,*h*,*i*]perylene	6-ring	<0.08	<0.08	<0.08	<0.08	2	2.8	1.9	1.2	0.99
Sum of PAHs[b]		785	996	1330	1430	468	672	1255	318	370
Sum of B[*a*]P equivalents[b,c]		0.16	0.23	1.06	0.05	4.7	8.52	10.25	4.17	2.40
Fluorescence (EU/g)		64.2	66.5	82.4	30.1	157	336	181	164	138

B[*a*]P, benzo[*a*]pyrene; EU, emission unit; LR, lab roofing; PAH, polycyclic aromatic hydrocarbon; TP, tank paving; TR, tank roofing.
[a]Results adapted from Kriech et al. (2007).
[b]One-half of the detection limit value was used to calculate total PAHs or B[*a*]P equivalents.
[c]Based on EPA (1993) relative potency factors (RPFs).

fumes were collected at 201 °C at a distance of 0.97 m (Kriech et al. 2007). The laboratory fumes were generated using methods similar to those used in a carcinogenicity bioassay (Sivak et al. 1997), using a temperature of 232 °C. PAH concentrations (total of 18 PAHs) from field-matched paving or roofing fumes were generally similar across the four source types (Table 5). However, the lab-generated roofing fumes (sample LR-A) contained significantly higher concentrations of PAHs than matched field-generated fumes (sample TR-A). For example, the total PAH concentration from the lab sample was 672 mg/kg, compared to 468 mg/kg from the field-matched sample (Table 5). These results are consistent with those presented in the "Asphalt chemistry and composition" section, in that PAH concentrations are generally similar from sample to sample and contain higher proportions of 2- to 3-ringed PAHs as compared to 4- to 6-ringed PAHs, and that emissions tend to increase with higher temperature (as evidenced when comparing field [TR-A] to lab [LR-A] samples in Table 5).

Occupational dermal exposure

Limited data are available (i.e. a few published studies for only a small subset of workers) for field-based dermal exposure estimates (van Rooij & Jongeneelen 2007). Three studies have examined dermal absorption during asphalt roofing (Hicks 1995; McClean et al. 2007; Wolff et al. 1989) and one study has examined this route of absorption from asphalt paving (McClean et al. 2004). Total PAH concentrations from skin wipe samples averaged 75 ng/cm^2 (pre-shift) and 876 ng/cm^2 (post-shift) and included exposure to coal tar during roof tear-off (Wolff et al. 1989). In contrast, Hicks (1995) collected post-shift dermal wipe samples from workers using BURA without coal tar exposure and found detectable levels of only 1 of 16 PAHs (naphthalene = 510–520 ng/cm^2). McClean et al. (2007) applied dermal patches (over multiple work days) to estimate skin exposures of 26 roofers to PAHs in asphalt fumes during hot application of BURA (Table 6). The results for individual PAHs (pyrene and B[*a*]P) were consistent with findings of previous studies, and on-roof workers were found to have higher dermal exposures. The removal of old coal tar pitch roof (tear-off) was associated with increased PAH exposures. Based on an evaluation of task-based exposures on non-coal tar pitch days, the best estimate of dermal exposures related to asphalt without coal tar were represented by the adjusted mean PAH concentrations of 340 ng/cm^2 for tear-off, 161 ng/cm^2 for put-down and 77 ng/cm^2 for kettlemen (McClean et al. 2007). Torassen et al. (2001) reported the air concentrations (TWA concentrations over a shift) for these roofing workers. Air concentrations for coal tar-exposed roofing workers ($n = 19$) ranged from 0.87 to 1.19 mg/m^3 TPM, 0.19–0.73 mg/m^3 BSF, and 95.5–198.2 μg/m^3 total PAHs. Air concentrations for non-coal tar-exposed roofing workers ($n = 7$) ranged from 0.24 to 0.31 mg/m^3 TPM, 0.08 to 0.18 mg/m^3 BSF and 16.4 to 33.9 μg/m^3 total PAHs. The authors concluded that an important pathway of exposure to PAHs during

Table 6. Dermal exposure (ng/cm^2) to PAHs during asphalt application.

	Total PAHs			Pyrene			B[*a*]P		
Worker type	*N*	GM	Range	*N*	GM	Range	*N*	GM	Range
Roofing activities[a]									
Roof workers[b]	71	898	48–34 014	71	11	<2.4–221	71	3.3	<0.5–59
Kettleman	19	299	40–4558	19	4.5	<2.4–34	18	0.9	<0.5–20
Tear-off[b]	41	886	49–33 538	41	11.5	<2.4–168	41	4.6	<0.5–84
Put-down	56	344	48–21 437	55	3.8	<2.4–150	54	1.0	<0.5–59
Paving activities[c]									
Paving	59	89	46–751	59	3.5	2.7–25	59	–[d]	0.8–2.5
Milling	39	–[d]	43–757	39	–[d]	5.3–7.1	39	–[d]	0.7–1.2
Roadside construction	11	45	45–246	11	–[d]	<2.6	11	–[d]	1.0–1.2

B[*a*]P, benzo[*a*]pyrene; GM, geometric mean; *N*, sample size; PAH, polycyclic aromatic hydrocarbon.
[a]Results adapted from McClean et al. (2007).
[b]Includes exposure to coal tar pitch during roof tear-off.
[c]Results adapted from McClean et al. (2004).
[d]GM not reported due to low level of detection.

roofing was via skin absorption, particularly when coal tar pitch roofing materials were present. In addition, alternative PAH sources that could have contributed to dermal exposure include smoking, construction materials, diesel exhaust or other ambient sources near the job sites.

The total dermal PAH concentrations reported for non-coal tar days (77–340 ng/cm^2) for asphalt roofing workers were similar to those reported for paving workers (also without coal tar exposure) (McClean et al. 2007). Total PAH concentrations in dermal patches from paving workers ranged from 43 to 757 ng/cm^2 (GM ranged from 45 to 89 ng/cm^2) (McClean et al. 2004). Total PAHs were also measured in air and ranged from 0.3 to 40 μg/m^3 (GM ranged from <0.2 to 4.1 μg/m^3).

A key component of dermal exposure assessment is evaluating the amount of chemical absorbed within and across the skin barrier. Few published data exist to characterize absorption rates for asphalt mixtures. However, recent dermal absorption studies of asphalt fume condensates (roofing and paving) were performed with an *in vitro* method that used human skin sections (Roy 2006a, b, c, d; Roy et al. 2007). These studies examined the same roofing asphalt fume condensates (TR-A, TR-B, TR-C, TR-D) generated by Kriech et al. (2007) (Table 5). The dermal study for paving asphalt fume condensate was performed using the same sample examined in a 2-year inhalation rodent bioassay (Fuhst et al. 2007; Roy et al. 2007). The experimental procedure examined human skin samples exposed to the asphalt samples under infinite dose conditions for 48 h. Tissue samples (250–320 μm thick) were mounted in a Franz diffusion unit with a 6% aqueous solution of polyoxyethylene 20 oleyl ether as the receptor fluid (to enhance solubilization of lipophilic materials) maintained at 37 °C (Roy 2006a, b, c, d; Roy et al. 2007). The dermal flux (i.e. quantitative measure of the rates of systemic uptake of chemicals via the dermal route) was measured using 200 μL of neat fume condensate per diffusion cell and reported for each of the samples over the 48-h experiment (Roy 2006a, b, c, d; Roy et al. 2007). The rates of dermal penetration (for total 3- to 6-ringed PAHs) in these experiments ranged from 5.0–120 ng/cm^2/h, and the estimated dermal flux of B[*a*]P in these samples ranged from 0.016 to 0.17 ng/cm^2/h (average from the five samples was 0.08 ng B[*a*]P/cm^2/h). By comparison, Moody et al. (2007) reported dermal flux for B[*a*]P of 10 ng/cm^2/h (in the receiver fluid) and 27 ng/cm^2/h (in the skin). Similar to the Roy et al. studies, Moody et al. (2007) also conducted an *in vitro* experiment (42 h) using human female breast skin and measured absorption of radiolabeled pure B[*a*]P in acetone. Further, these *in vitro* experiments demonstrate that the dermal flux of B[*a*]P in asphalt mixtures is reduced relative to a pure B[*a*]P solution (i.e. ratio of B[*a*]P in asphalt vs. acetone is 0.08/10 ng/cm^2/h = 0.8%). These studies indicate that a limited amount of PAHs (particularly B[*a*]P) in asphalt condensates is absorbed across the skin barrier and available to systemic circulation (or alternatively, a large portion does not penetrate the stratum corneum over 48 h, the duration of the experiment).

The *in vitro* dermal studies examined penetration through the skin and hence speak directly to systemic exposure via the dermal route. There is little information, however, on absorption into and retention within the dermal tissue itself, which is a key parameter for skin cancer dosimetry. Storm et al. (1990) examined dermal absorption and metabolic activity within rodent and human skin. Approximately 60% of the dermal dose of B[*a*]P was absorbed in rodents compared to 31% in human skin, with most of this amount deposited within the skin (for human skin= 1% in receptor fluid and 30% in the skin depot). Storm et al.

(1990) noted that the absorption rate correlated with metabolic activity; metabolic activity (a key event in PAH carcinogenicity) was reported to be approximately 14 times greater in mouse skin than in human skin (i.e. aryl hydrocarbon hydroxylase [AHH] was 3.35 pmol/min-mg-protein in mice vs. 0.24 pmol/min-mg-protein in human skin) (Storm et al. 1990). Kao et al. (1985) also reported that *in vitro* 24-h percutaneous absorption of B[*a*]P was greater in mouse (10%) back skin (0.34 mm thick) versus human (<3%) leg skin (1.32 mm thick) samples. Moody et al. (2007) also examined B[*a*]P absorption in human skin. Total skin absorption (skin depot and receptor fluid) was 56.4% (after 24 h) and 49.7% (after 48 h), with most in the skin depot (45.4% and 39% after 24 and 48 h, respectively) and a lower proportion in the receptor fluid (10–11%). Approximately 30% of the B[*a*]P remained unabsorbed and was washed off the skin surface. Thus, these studies suggest that approximately 30–45% of dermally applied B[*a*]P is retained in the skin after 24–48 h of exposure. If we conservatively assume that B[*a*]P absorption and BURA absorption are generally similar (although the rate of absorption is likely much slower, as discussed above), we would expect 30–45% of the PAHs in BURA to be retained in the skin, and a fraction of that amount to be metabolized (e.g. 7% based on differences between metabolism in mice and humans). Considering both absorption and metabolism (e.g. 30–45% multiplied by 7%), approximately 2–3% of the B[*a*]P from these studies would be retained in human skin, metabolized and available to react with skin cells (keratinocytes). This relative absorption rate is likely overestimated for asphalt mixtures due to the small proportion of carcinogenic PAHs (4- to 6-ringed PAHs) and the slower rate of absorption.

Environmental monitoring information for inhalation and dermal exposure to BURA is sufficient for characterizing human exposure either to the heated asphalt mixtures or to several individual PAHs known to be biologically active. Personal-air sampling measurements from numerous studies of roofing workers have reported similar ranges of concentrations, and these ranges were typically below or near the current TLV of 0.5 mg/m^3 BSF (Table 4). Further, confounding factors have been identified in a number of instances, so that exposure to BURA can be distinguished from other potential contaminant sources. Similarly, dermal exposure concentrations (and limited information on absorption rates) are available for typical exposure conditions. Therefore, existing occupational exposure information is sufficient for characterizing both whole mixtures and component-based mixtures for BURA and are useful for conducting a quantitative risk assessment.

Epidemiology data

Epidemiology studies with sufficient concentration–response data for the exposure route of interest, when available, are best for assessing the health effects of chemical mixtures (EPA 2000). Cancer risks from asphalt exposures have been assessed in a number of epidemiology studies (reviewed by Fayerweather 2007; IARC 2013a; NIOSH 2000; Partanen & Boffetta 1994). IARC (2013a) noted that there was limited evidence for upper aerodigestive tract cancers (buccal cavity, pharynx, oesophagus and larynx) following occupational exposure to bitumen and bitumen emissions. Further, in a recent case–control study, Fogleman et al. (2015) reported no association between occupational asphalt exposure and head and neck squamous cell carcinoma. Below, we review the studies that evaluated occupational exposure to asphalt emissions during roofing applications (both in the USA and Europe) and lung and skin cancer risks. We focus on these cancer endpoints because we have analyzable data and can integrate information across multiple studies, including animal bioassays. We specifically focus on exposure conditions and potential confounding factors. The epidemiology assessment provides additional context for the quantitative cancer risk estimates.

Review of lung cancer data

Numerous cohort, case–control, cross-sectional, and proportionate mortality studies have evaluated lung cancer risk associated with exposure to asphalts and asphalt emissions (summarized in Tables 7 and 8). Lung cancer risk estimates (and 95% confidence intervals [95% CI]) are typically provided as relative risks (RRs), odds ratios (ORs), proportionate mortality ratios (PMRs), standard mortality ratios (SMRs), and standard incidence ratios (SIRs). The more credible studies reported RR estimates between 1 and 1.78; overall, however, the epidemiology data provide limited information regarding the association between exposure to asphalt emissions and lung cancer, primarily owing to study design limitations, such as exposure measurement error, and many confounding variables encountered in these occupational settings, such as coal tar, asbestos, and smoking (Fayerweather 2007; IARC 2013a; NIOSH 2000; Partanen & Boffetta 1994).

Cohort studies

Seven cohort studies have focused on asphalt workers. The largest study (79 822 workers) is an evaluation of the European multi-center cohort (also called the IARC

Table 7. Epidemiology studies of asphalt roofers and lung cancer.

Reference	Country	Study period	Subjects (*N*)		Roofer cases	Risk estimate (95% CI)
Cohort studies						
Boffetta et al. (2001, 2003a)	Denmark, Finland, Sweden	1953–2000	29 820		14	SMR = 1.33 (0.73–2.33)[a]
					14	RR = 1.34 (0.71–2.53)
Engholm et al. (1991)[b]	Sweden	1971–1985	3456		3	SMR = 2.79 (0.58–8.16)
					4	SIR = 3.62 (0.99–9.27)
Hammond et al. (1976)	USA	1960–1971	5339		22	SMR = 0.92 (0.57–1.39)[c]
					99	SMR = 1.59 (1.29–1.94)[d]
Hansen (1991)	Denmark	1959–1986	679		25	SMR = 2.90 (1.88–4.29)
Hrubec et al. (1992)[b]	USA	1954–1980	52		4	RR = 3.0 (1.30–6.75)
Menck and Henderson (1976)[b]	USA	1968–1970; 1972–1973	2000		11	RR = 4.96 ($p<0.05$)
Swaen and Slagen (1997)	Netherlands	1947–1980	866		39	SMR = 1.31 (0.93–1.80)
Case–control studies			Cases	Controls		
Morabia et al. (1992)[b]	USA	1980–1989	1793	3228	7	OR = 2.1 (0.7–6.2)
Schoenberg et al. (1987)[b]	USA	1967–1976	763	900	12	OR = 1.7 (0.68–4.4)
Vineis et al. (1988)	USA	1974–1981	2973	3210	45	OR = 1.4 (0.9–2.3)
Zahm et al. (1989)	USA	1980–1985	4431	11 326	6	OR = 2.1 (0.6–8.2)
Watkins et al. (2002)	USA	1977–1979	39	133	12	OR = 1.6 (0.6–4.6)
Richiardi et al. (2004)	Italy	1990–1992	1171	1553	9	OR = 2.0 (0.6–6.5)
McClean et al. (2011)[b]	USA	1998–2003	422	894	32	OR = 1.2 (0.7–2.1)[e]
						OR = 1.11 (1.01–1.22)[f]
Cross-sectional and proportionate mortality studies						
Dong et al. (1995)[b]	UK	1975–1987	15 007		26	PMR = 1.15 (0.67–1.96)
Leigh (1996)[b]	USA	1979–1981	12 545		41	RR = 3.29 (2.52–4.53)
Menck and Henderson (1976)[b]	USA	1968–1970	3938		11	RR = 4.96 (2.48–8.87)
Milham (1997)[b]	USA	1950–1989	676 161		86	PMR = 1.44 ($p<0.01$)
Stern et al. (2000)[b]	USA	1950–1996	11 370		778	PMR = 1.39 (1.31–1.48)
Wang et al. (1999)[b]	USA	1988–1994	NR		NR	PMR = 1.10 (0.99–1.22)

95% CI, 95% confidence interval; IARC, International Agency for Research on Cancer; *N*, sample size; NR, not reported; OR, odds ratio; PMR, proportionate mortality ratio; RR, relative risk; SIR, standard incidence ratio; SMR, standard mortality ratio.
[a]Adjusted for coal tar exposure.
[b]As cited by IARC (2013a) and Fayerweather (2007).
[c]Exposed for 9–19 years.
[d]Exposed for 20 years.
[e]Risk estimate for workers ever exposed.
[f]Risk estimate for cumulative exposure.

cohort), which consists of asphalt workers followed from 1953 to 2000 and uses a job-exposure matrix to estimate exposures (Bofetta et al. 2001, 2003a, b; Burstyn et al. 2000, 2003). Subjects were identified from company records and included individuals who worked for at least one season. Exposure was characterized as "ever" or "never" exposed to asphalts and/or coal tar based on company records. Lung cancer risk among roofers and waterproofers was not significantly elevated (14 deaths, SMR = 1.33, 95% CI = 0.73–2.23). There was no significant heterogeneity among lung cancer SMRs between countries ($p = 0.4$), although a large variation in the SMRs between countries was noted for building- and ground-construction workers who were used as unexposed referents in the internal analysis ($p = 0.01$) (IARC 2013a).

Burstyn et al. (2000, 2003) conducted further analyses of this cohort, including more detailed industrial hygiene measurements. Exposure estimates from primarily paving operations were applied to all asphalt workers. GM and arithmetic mean (AM) concentrations for all paving operations were reported for asphalt fumes (GM = 0.28 mg/m^3, AM = 1.91 mg/m^3), asphalt vapor (GM = 1.86 mg/m^3, AM = 7.59 mg/m^3), and B[*a*]P (GM = 8.58 ng/m^3, AM = 95.8 ng/m^3). Mean concentrations for mastic asphalt operations were also reported for asphalt fumes (GM = 2.29 mg/m^3, AM = 13.4 mg/m^3), asphalt vapor (GM = 2.06 mg/m^3, AM = 4.14 mg/m^3), and B[*a*]P (GM = 61.6 ng/m^3, AM = 715 ng/m^3). Lung cancer risks for those workers ever or never exposed to asphalt were similar (SMR = 1.08, 95% CI: 0.99–1.18, vs. SMR = 1.05, 95% CI = 0.92–1.19, respectively) (Bofetta et al. 2003b). In the "coal tar free" subcohort, lung cancer risk was significantly elevated (SMR = 1.23, 95% CI: 1.02–1.48), but no significant associations were noted with exposure duration, cumulative exposure or average exposure to asphalt (Bofetta et al. 2003b). Olsson et al. (2010) performed a follow-up nested case–control study, further described below.

Hammond et al. (1976) evaluated mortality in a historical cohort of 5939 roofers and waterproofers exposed to coal tar pitch and asphalt (study period 1960–1971). Duration of union membership was used as a surrogate for exposure duration. As reported by IARC (2013a), lung cancer risk was elevated among roofers in this cohort with over 20 years of union membership (SMR = 1.59, 95% CI: 1.29–1.94). The authors did not

Table 8. Lung cancer risks associated with inhalation and dermal exposure to asphalt fumes and condensates in the IARC cohort.

Inhalation exposure to asphalt fumes		Dermal exposure to asphalt fume condensates	
Exposure category	OR (95% CI)[a,b]	Exposure category	OR (95% CI)[a,b]
Never exposed	1	Never exposed	1
Ever exposed	1.12 (0.84–1.49)	Ever exposed	1.17 (0.88–1.56)
Exposure duration (year)			
0.33–7.99	1.19 (0.84–1.69)	0.33–7.99	1.22 (0.86–1.74)
8.00–15.49	1.26 (0.87–1.83)	8.00–15.49	1.34 (0.93–1.94)
15.50–25.99	1.23 (0.94–1.79)	15.50–25.99	1.35 (0.93–1.96)
26.00–54.00	0.74 (0.49–1.11)	26.00–54.00	0.72 (0.47–1.10)
	Linear trend p values (0.37)		Linear trend p values (0.50)
Cumulative exposure (unit years)			
0.18–9.55	1.31 (0.93–1.85)	0.59–61.54	1.21 (0.85–1.72)
9.56–28.17	0.99 (0.68–1.45)	61.55–185.25	1.22 (0.84–1.76)
28.18–68.00	1.16 (0.78–1.72)	185.26–407.07	0.99 (0.66–1.49)
68.01–620.48	0.77 (0.50–1.19)	407.08–4003.76	1.21 (0.79–1.84)
	Linear trend p values (0.39)		Linear trend p values (0.58)
Average exposure (units)			
0.08–0.97	1.20 (0.84–1.71)	0.29–6.62	1.10 (0.77–1.57)
0.98–2.20	1.15 (0.78–1.70)	6.63–13.44	1.21 (0.83–1.76)
2.21–3.61	0.90 (0.60–1.34)	13.45–23.06	1.25 (0.84–1.87)
3.62–16.67	1.16 (0.78–1.73)	23.07–94.11	1.23 (0.81–1.88)
	Linear trend p values (0.80)		Linear trend p values (0.26)

95% CI, 95% confidence interval; IARC, International Agency for Research on Cancer; OR, odds ratio.
[a]As reported in the nested-case control study of a European Multi-center Cohort by Olsson et al. (2010).
[b]Adjusted for sex, country, age, tobacco pack-years and coal tar exposure.

account for smoking or exposure to coal tar, so which specific exposures contributed to increased risk, or to what extent, cannot be determined (IARC 2013a).

Engholm et al. (1991), as summarized by IARC (2013a), reported cancer mortality risks in Swedish asphalt workers (from 1971 to 1985, with an average 11.5 year follow-up). This cohort was also included in the IARC multi-center study. Excess lung cancer risk was reported for roofers (SMR = 2.79 and SIR = 3.62, based on four cases and three deaths), but CIs and p values were not reported. IARC (2013a) estimated CIs (Table 7) that indicate these risks were not statistically significant. The authors did not report information for other potentially confounding chemical exposures (e.g. coal tar).

Hansen (1991) followed a Danish cohort of mastic asphalt workers with death records from 1959 through 1986. The authors suggested that mastic workers experience similar exposure to asphalt fumes as that experienced by roofers, although data were not provided to support that suggestion. In fact, studies that monitored both roofing and mastic workers with similar exposure methods (Brandt et al. 1985, Ruhl et al. 2006) indicate that mastic workers are exposed to much higher asphalt fume levels than roofers. The Hansen cohort included workers (3748 of 7434 person-years) that may have been exposed to coal tar during work prior to 1930. Among mastic workers aged 40–89 years, a significantly increased risk of lung cancer was observed, even when urbanization and smoking were considered (SMR = 2.24, 95% CI: 1.45–3.30). The reported concentration of asphalt fume condensate in 35 personal samples taken during flooring ranged from 0.5 to 260 mg/m^3 (median of 19.7 mg/m^3) (Hansen 1991). This study has been criticized for possible selection bias, confounding by alcohol consumption, residual confounding by tobacco smoking and urbanization, and confounding by the presence of coal tar (see NIOSH 2000; Partanen & Boffetta 1994; Wong et al. 1992). NIOSH (2000) indicated that the major criticisms include the lack of control for confounding by smoking and urbanization, possible confounding by coal tar, biases in the selection of the cohort, and inadequate data on work and exposure histories. The results of all of these studies have not been reconciled, and, therefore, how much of the estimated risk is attributable to coal tar or other factors remains unclear.

Hrubec et al. (1992), as summarized by IARC (2013a), evaluated the mortality of roofers and waterproofers in a cohort of US veterans (study period 1954–1980). Mortality from respiratory cancer was slightly elevated (RR = 3.0, 95% CI: 1.30–6.75) based on four cases. Although this study controlled for smoking, IARC (2013a) noted that the sample size was limited for roofers, and that subjects likely had been co-exposed to coal tar and/or asbestos.

An increase in lung cancer mortality (SMR = 4.96, 95% CI: 2.5–8.9) among roofers working in Los Angeles County, California was reported by Menck & Henderson (1976). Results were based on last occupation from death certificates (from 1968 to 1970 and 1972 to 1973) and length of employment reported at admission or by next of kin. No information was provided for

confounding factors such as exposure to coal tar, exposure to asbestos or smoking history (IARC 2013a).

Swaen & Slangen (1997) performed a retrospective cohort study in a population of 866 roofers employed between 1947 and 1980 and followed this population through 1988. Lung and trachea cancer risk were not significantly elevated in this cohort (SMR = 1.31, 95% CI: 0.93–1.80). The statistical analysis did not control for exposure to cigarette smoke or coal tar.

Case–control studies

Several case–control studies have shown evaluated cancer risks in asphalt workers (Table 7). Olsson et al. (2010) performed a nested case–control study of European asphalt workers included in the IARC multicenter cohort study (described above). The objective of this study was to assess the contribution of asphalt and other occupational exposures to the risk of lung cancer among all asphalt workers (including pavers, roofers, and other job categories). A total of 433 workers who died between 1980 and the end of follow-up (between 2002 and 2005) and 1253 controls matched to these cases based on year of birth and country were included from the original cohort (note that not all cases identified were included, as only 65% agreed to be interviewed). Both inhalation and dermal exposure were assessed semi-quantitatively using historic industrial hygiene measurements (Olsson et al. 2010). Cases included workers with at least two seasons of employment who were diagnosed with or died from lung cancer between 1980 and 2005. Risks were not significantly elevated for those ever exposed to asphalt fume (inhalation OR = 1.12, 95% CI: 0.84–1.49) or ever exposed to asphalt condensate (dermal OR = 1.17, 95% CI: 0.88–1.56). No significant trends were observed for exposure duration, average exposure or cumulative exposure to asphalt fumes or condensates (see Table 8). Although roofers were not assessed separately, excluding this group had no significant effect on the risk estimates (Olsson et al. 2010). The authors reported: (a) no significant association was found between indicators of inhalation and dermal exposure to bitumen and lung cancer, (b) other known or suspected occupational lung carcinogens in the asphalt industry had no detectable effect, with the possible exception of coal tar and (c) prevalence of tobacco smoking was higher in the study population than in the general population, as indicated by national surveys, a finding that might have biased the results of the cohort study away from the null (Olsson et al. 2010).

Morabia et al. (1992) conducted a case–control study of lung cancer and occupation in the USA between 1980 and 1989. Cases (1793 men) were matched by age, race, and smoking history to 3228 controls (men) with other diagnoses. Lung cancer risks were not elevated among roofers or slaters (OR = 2.1, 95% CI: 0.7–6.2). This estimate was not adjusted for asbestos or coal dust exposure, and it would likely have been lower if it had been.

In a population-based case–control study, Schoenberg et al. (1987), as cited by IARC (2013a) examined associations between lung cancer and occupation in roofers and slaters from New Jersey, USA. The cases ($n = 763$) were white male residents diagnosed with lung cancer between September 1980 and October 1981. The controls ($n = 900$) were sampled at random from driver's license records or death certificate registers and matched on age, race, area of residence and date of death. Occupational history was obtained from next of kin. A smoking-adjusted OR of 1.7 (95% CI: 0.68–4.4) was reported for the occupational category of roofers and slaters. No information was provided regarding exposure to asphalts or other chemicals.

Vineis et al. (1988) pooled data from five case–control studies of cancer (conducted during the 1970s and 1980s), including the study by Schoenberg et al. (1987, as cited by IARC 2013a). The pooled data set included 2973 male cases and 3210 controls. Risk from exposure to asphalt was not specifically assessed, but a smoking-adjusted OR of 1.4 (95% CI: 0.9–2.3) was reported for lung cancer among roofers and asphalt workers.

Zahm et al. (1989) studied associations between occupation and lung cancer using a cancer registry in Missouri, USA. The cases (all white men) were diagnosed between 1980 and 1985, and controls had other non-smoking related cancers diagnosed during the same period. The OR for lung cancer in roofers was 2.1 (95% CI: 0.6–8.2), based on six exposed cases and seven controls, and the OR did not differ significantly by histological subtype. The main limitations of this study are that it used job titles as indicators of exposures and that it did not assess chemical exposures.

Three population-based case–control studies were performed that evaluated workers occupationally exposed to asphalt (Table 7). Richiardi et al. (2004) evaluated 1171 cases and 1553 controls and reported an OR of 2.0 (95% CI: 0.6–6.5) for a construction category that included roofers and asphalt workers. McClean et al. (2011), as cited by IARC (2013a) examined 422 cancer cases and 894 controls in San Francisco, California, and reported an OR of 1.2 (95% CI: 0.7–2.1) for those ever exposed to asphalt and tar. Finally, Watkins et al. (2002) examined roofing asphalt manufacturing workers between 1977 and 1997. Thirty-nine cancer cases and 133 controls were matched on age, race and year

of birth. The OR for those ever exposed to asphalt fume was 1.59 (95% CI: 0.60–4.57); exposed for <20 years was 2.27 (95% CI: 0.74–7.73); and exposed for 20 years was 1.06 (95% CI: 0.30–3.65). No dose–response relationship was reported, and lung cancer was strongly correlated with cigarette smoking.

Cross-sectional and proportionate mortality studies

IARC (2013a) and Fayerweather (2007) reviewed several cross-sectional and proportionate mortality studies of populations exposed to asphalt (Table 7). PMRs ranged from 1.10 to 1.44, and RRs ranged from 3.29 to 4.96. Although some of the risk estimates were relatively strong and statistically significant, none of these studies conducted extensive exposure analyses or collected information on co-exposures to coal tar or asbestos, or cigarette smoking (IARC 2013a). PMR studies are generally less reliable than cohort and case–control studies because PMRs tend to overestimate mortality incidence (Wong 1983).

Meta-analyses

Meta-analyses of asphalt/lung cancer studies were performed by Partanen & Boffetta (1994) and Fayerweather (2007). Partanen and Bofetta (1994) included 20 studies published through 1993 and covering the period from 1945 to 1989. Meta-RRs for lung cancer were 0.87 (95% CI: 0.76–1.08) among pavers and highway-maintenance workers and 1.78 (95% CI: 1.50–2.10) among roofers. Four of the 20 studies adjusted for smoking, and the aggregated RR for lung cancer was 2.0 (95% CI: 1.3–2.8) based on those four studies. Partanen & Boffetta (1994) were not able to differentiate the effects of coal tar exposure from exposure to asphalt emissions.

Fayerweather (2007) conducted a meta-analysis of 27 peer-reviewed epidemiology studies. Unlike Partanen & Boffetta (1994), Fayerweather (2007) controlled for possible confounding from exposure to coal tar by assuming concentrations of B[*a*]P of 20 $\mu g/m^3$ for coal tar roofing and 10 $\mu g/m^3$ for coal tar paving. Although initial associations were statistically significant, after adjustment, the meta-RRs dropped from 1.67 (95% CI: 1.39–2.02) to 1.10 (95% CI: 0.91–1.33) for roofers, and from 0.98 (95% CI: 0.81–1.18) to 0.96 (95% CI: 0.80–1.16) for pavers.

Mundt et al. (2007) performed a meta-analysis to evaluate commonly recognized confounders (i.e. smoking, coal tar and asbestos exposure) of lung cancer risk among roofers. Based on an external confounder adjustment method (Axelson & Steenland et al. 1988, as cited by Mundt et al. 2007), the authors examined the RRs from smoking (RR = 1.17–1.52), asbestos (RR = 1.36–1.78), and coal tar (RR = 1.04–2.32). The results indicate that cancer risk estimates for roofers fall within the range of risks from confounding variables, and that existing epidemiology studies cannot be used to conclude reliably that asphalt exposure alone results in an increased risk of lung cancer.

Armstrong et al. (2004) conducted a meta-analysis of studies of industries with occupational exposures to PAH mixtures. They reviewed three asphalt epidemiology studies by Hammond et al. (1976), Hansen (1991) and Swaen and Slangen (1997). Unit relative risks (URRs) were estimated per 100 $\mu g/m^3$ years of cumulative B[*a*]P exposure for each study. The mean URR for asphalt workers was 17.5 (95% CI: 4.21–72.78) at 100 $\mu g/m^3$ B[*a*]P years. Armstrong et al. (2004) did not include the IARC multi-center study in the meta-analysis but provided a point URR for the IARC cohort study of 44.9 (95% CI: 25.0–64.8). The results for asphalt workers were noted to be higher than for other industries but imprecisely estimated. The results of this study are uncertain due to the wide variability in URR estimates, as well as the exposure estimates and confounding variables for each of the individual studies as described previously.

Bosetti et al. (2007) also performed a limited meta-analysis of four cohorts: Hammond et al. (1976), Hansen (1991), Swaen and Slangen (1997) and Boffetta et al. (2003a). The pooled RRs for lung cancer were 1.51 (95% CI: 1.28–1.78) for roofers and 1.14 (95% CI: 1.07–1.22) for asphalt workers. The authors suggested that a modest increase in lung cancer was observed for roofers, but the results may have been affected by bias or confounding factors.

Review of skin cancer data

Few epidemiology studies have assessed skin cancer risks in roofers. The IARC multi-center cohort study (Boffetta et al. 2003b) did not report a significant association between asphalt exposure (all workers) and non-melanoma skin cancers (SMR = 0.74, 95% CI: 0.41–1.21). All other studies of skin cancer reported risks that were not statistically significant (and both below and above unity). Hammond et al. (1976) reported elevated SMRs (but no *p* values) by exposure duration (SMR = 4.65 for 19 years and SMR = 4.0 for >20 years of union membership). As mentioned in the "Review of lung cancer data" section, this study did not assess exposure to asphalt and did not correct for coal tar or smoking. Povarov et al. (1988) evaluated 1486 subjects from Estonia (exposed 1974–1984) and reported a non-significant incidence rate ratio of 1.5 (four cases). Hansen (1989) reported no elevated risks (SIR = 0.67,

95% CI: 0.14–1.96) for 679 workers (three cases) from Denmark (exposed 1959–1984). Milham (1997) and Stern et al. (2000) performed proportional mortality studies (PMR = 1.0 and 0.69, respectively) and reported no significant increase in skin cancer mortality.

Partanen and Bofetta (1994) performed a meta-analysis of non-melanoma skin cancer risks and reported an increased risk for all asphalt workers (meta-RR = 1.74, 95% CI: 1.07–2.65) and roofers (RR = 4.0, 95% CI: 0.83–11.7) but also did not control for exposure to coal tar and relied only on the Hammond et al. (1976) study to evaluate risks in roofers.

Evaluation of epidemiology evidence

In our analysis of human data on risks of cancers, we found that the most credible studies estimated an RR of lung cancer among asphalt roofers of between 1.1 and 1.78, and no studies reported a statistically significant risk of skin cancer. To evaluate whether estimated lung cancer risks were indicative of causation, we considered the strength of associations, exposure assessment, multiple comparisons, confounding and bias, exposure-response and consistency of results.

Strength of association: Generally, the larger the RR, the more likely the observed association is causal. Small magnitudes of association (e.g. RRs 2.0) are less likely to represent causal associations and may be a result of bias (from uncontrolled confounding, for example). The preponderance of the RRs from cohort and case–control studies (Table 7) suggests weak associations between inhalation of asphalt fumes and lung cancer (i.e. risk estimates were mostly below 2.0, and 95% CIs mostly included unity).

Exposure assessment: Investigators typically estimated exposure from limited historic quantitative measurements or job category. Although a few studies (notably, the IARC multi-center cohort studies) attempted to conduct semi-quantitative estimates of exposure, most did not provide rigorous estimates of airborne or dermal exposure concentrations. The lack of precise exposure data in each epidemiology study likely led to exposure measurement error and/or exposure misclassification, which could have biased the results (Fayerweather 2007; IARC 2013a; NIOSH 2000; Partanen & Boffetta 1994).

Multiple comparisons: Many studies evaluated many outcomes several ways (e.g. by looking at exposure different ways or conducting stratified analyses). Unless corrections are made for multiple comparisons, finding statistical significance in one or a few comparisons may occur by chance even when no true effect exists (dos Santos Silva 1999). Early cohort, case–control, and cross-sectional studies focusing on asphalt workers seldom corrected for multiple comparisons. Because of this omission, some of the reported increased risks are not likely indicative of causation.

Confounding and bias: In many occupational studies, asphalt exposure occurred concurrently with co-exposures to a number of factors. Although some studies attempted to control for one or more of these confounding factors statistically, only the nested-case–control study by Olsson et al. (2010) provided a robust analysis to separate asphalt exposure from exposures to other chemicals. They found no significant increase in lung or skin cancer risk after controlling for cigarette smoke and coal tar (the primary confounders recognized for asphalt workers). Fayerweather (2007) and Mundt et al. (2007) provided additional analyses that corroborate these results and indicate that controlling for these confounders attenuates cancer risks. Thus, historic epidemiology investigations provide limited evidence to reliably assess cancer risks from asphalt exposures among roofing workers.

Exposure-response: If exposure to asphalt fumes or condensates were a causal factor for lung and/or skin cancer, one would expect cancer risks to increase with exposure both within and among studies. Unfortunately, few studies performed assessments of exposure-response relationships. Olsson et al. (2010) provided the most robust assessment of exposure by duration and intensity while controlling for confounders and reported no significant trends between lung cancer risk and increasing exposures (Table 8). These analyses were performed for all types of asphalt workers, but when roofers were removed from the analyses, no changes in risk estimates were observed.

Consistency of results: Both Fayerweather (2007) and Partanen and Boffetta (1994) assessed the heterogeneity among epidemiology investigations and concluded that a high level of variability exists between the risk estimates. They concluded that this variability might be partially attributable to the lack of controls for confounding factors, as analyses that adjusted for confounders had reduced estimates of heterogeneity (Fayerweather 2007). Overall, the results of historic and recent epidemiology studies are not consistent, perhaps owing to the varied types of analyses, subject groups, and confounding variables.

Based on this evaluation, we conclude that the weight of epidemiology evidence does not support roofing asphalt as a risk factor for lung cancer. Further, the data available for developing quantitative estimates of risk are limited, because the exposure measures and confounding factors are imprecise. Of particular concern is that many of the risks attributed to asphalt may instead be attributable to coal tar, asbestos, smoking or another factor, and determining what proportion of the risk, if any, is attributable to asphalt is not possible. Overall, the epidemiology data do not provide a clear basis for estimating lung or skin cancer risk (nor other types of cancer risks). However, because results from experimental animal studies describe skin and lung tumors following exposure to asphalt emissions, we can assess those findings in relation to the results from the human experience.

Experimental animal data

Chronic animal bioassays to evaluate the carcinogenicity of fumes or fume condensates from heated asphalt mixtures have been performed using a variety of test species and exposure regimes (Arthur D. Little, Inc. 1981, 1989; Clark et al. 2011; Fuhst et al. 2007; IARC 2013a; Niemeier et al. 1988; Nesnow et al. 1983; Sivak et al. 1997; Thayer et al. 1983). From the available toxicological literature, we identified the most appropriate studies for use in a quantitative assessment of the dose–response relationship between exposure to BURA fume and the development of cancer following dermal or inhalation exposures (Table 9). The available animal data provide quantitative tumor response data for specific carcinogenic endpoints (i.e. kin tumors from skin painting or lung tumors from inhalation). Thus, we used the available animal results and their indication of risk and extrapolated to humans. Owing to limitations with quantifiable human exposure-response data (see the "Epidemiology data" section), we analyzed the animal studies by the exposure routes and endpoints available. In addition, these studies allowed us to examine key parameters, such as whole-mixture exposure-response, mixture-component exposure-response, comparisons of potency and component interactions, necessary to investigate chemical mixtures per EPA (2000) guidance.

Inhalation bioassays

One animal bioassay has been performed using a paving asphalt mixture (Fuhst et al. 2007). The study design consisted of a nose-only exposure to asphalt fumes (particulate and vapor) generated from condensate samples that were collected from the headspace of hot asphalt storage tanks using a protocol shown to yield fumes with compositions that match worker exposures during representative paving operations. Male and female Wistar rats ($n = 50$ and 86, respectively) were exposed to clean air, or to 6.8, 34.4 or 172 mg/m^3 total hydrocarbon for 6 h/day, 5 days/week, for 24 months. Tumor incidence data for each treatment group were presented. No significant increase in the number of tumor-bearing animals (based on comprehensive histopathological analysis) was observed in the exposed groups compared to concurrent controls. In addition, no treatment-related significant increases were identified in the incidence of organ-related tumors. A single animal in the highest dose group had a poorly differentiated adenocarcinoma in the nasal cavity. No data was available on the time to tumor. The results of this study are essentially negative, with no clear elevation in incidence of any tumor among exposed rats (i.e. no statistically significant increase in tumor-bearing animals or any increase in organ-specific tumors). Owing to these shortcomings, these data are too limited to assess a multi-dose–response relationship. However, the data are useful for identifying a plausible upper bound for our inhalation cancer assessment (i.e. inclusive of lung and any other inhalation related cancer), as discussed in the following sections.

Dermal bioassays

Several bioassays have been performed using dermal application of asphalt to mice (Table 9). These data provide an opportunity to examine the potency of BURA fumes generated by different methods, notably at different temperatures and by methods that capture field fumes rather than creating fumes in a laboratory. It also enables examination of the consistency of findings over several determinations.

NIOSH initially conducted a skin painting bioassay to assess the carcinogenic potential of condensed volatiles from roofing asphalts under differing temperatures and light conditions (Arthur D. Little, Inc. 1981; Niemeier et al. 1988). In this study, CD1 and C3H/HeJ male mice (test groups of 50 animals, ages 12–15 weeks) were exposed to laboratory-generated condensed fumes from Type I or Type III asphalts (produced by distillation and air blowing of Arabian crude). Asphalt fumes were collected at temperatures of either 232 or 316 °C. The condensed fumes were diluted in 1/1 cyclohexane/acetone applied to the clipped interscapular area every 2 weeks with 50 μL of the appropriate test material and under two light conditions (with and without simulated sunlight). Tumor incidence was monitored for 18 months.

Table 9. Summary of rodent carcinogenicity bioassays for BURA and paving fumes.

Reference	Route	Test species	Exposure method	Exposure duration	Test doses	Results
Fuhst et al. (2007)	Inhalation	Wister rat (50 males or 86 females per group)	Nose-only exposure to lab-generated bitumen fumes (particulate and vapor) from a paving asphalt	6 h/day, 5 days/week, for 24 months	0, 6.8, 34.4 or 172 mg/m^3 (total hydrocarbon)	No significant increase in number of tumor-bearing animals compared to controls
Arthur D. Little, Inc. (1981); Niemeier et al. (1988)	Dermal	CD-1 or C3H/HeJ male mice (50 per group)	Type I or Type III asphalts (collected at 232 or 316 °C) applied to the clipped interscapular area	Biweekly application for 78 weeks	25 mg of total solid per application and a positive control group was included (0.01% B[*a*]P/50 μL)	Average tumor latency period ranged from 39.5 to 56.1 weeks among the C3H/HeJ groups, and from 47.4 to 76.5 weeks among the CD-1 mice
Arthur D. Little (1989) and Sivak et al. (1997)	Dermal	Male C3H/HeJ mice (30/group)	Lab-generated BURA Type III "steep" asphalt (at 316 °C) applied to the interscapular area	Biweekly application for 104 weeks	Asphalt fume condensates were fractionated in five different fractions (A–E). A positive control group was included (0.05, 0.5 and 5 μg B[*a*]P/50 μL application). Forty-two combinations of raw asphalt, asphalt fractions, or B[*a*]P plus asphalt fractions were evaluated	All combinations including fractions B or C induced tumors. Combinations without fractions B or C failed to induce any tumors
Clark et al. (2011) and MPI (2010)	Dermal	Male C3H/HeNCrl mice (80 treated and 50 controls)	Field and lab-generated BURA Type III asphalt and field generated paving asphalt applied to the clipped back (10% of total surface area)	Biweekly application for 104 weeks	Fume condensates applied in a volume of 37.5 μL (25 mg) mineral oil. A positive control group was included (0.05% B[*a*]P/37.5 μL toluene)	Skin carcinomas were induced in treated animals: paving asphalt (0%), BURA field fume (13%), BURA lab fume (55%), and B[*a*]P treatment (69%)

B[*a*]P, benzo[*a*]pyrene; BURA, built-up roofing asphalt.

In addition, two coal tar pitch condensates were also examined along with a positive control (i.e. B[*a*]P). The authors concluded from this study that asphalt and coal tar fume condensates were carcinogenic, the non-pigmented CD-1 mouse strain was less sensitive (i.e. lower tumor incidence), increasing temperature increased tumorigenic response (in C3H/HeJ mice) and simulated sunlight inhibited the tumorigenic response (in C3H/HeJ mice). Single dose levels were examined in this experiment, and individual animal data were not reported, limiting the utility of these data for quantitative dose–response or time-to-tumor modeling. The existing data are useful, however, for examining the relative potency of lab-generated asphalt fumes to that of B[*a*]P and for comparison to subsequent skin painting bioassays.

NIOSH conducted a second chronic skin-painting study to evaluate laboratory-generated whole fume Type III BURA condensates. They also fractionated the total fume condensate into five chemical fractions (A–E) and tested the individual fractions, alone and in various combinations, for dermal carcinogenicity (Arthur D. Little, Inc. 1989; Sivak et al. 1997). Each material was tested at only one dose, and the doses for the fractions constituted amounts intended to be equal to their representation in the whole-fume exposure. The incidence of skin tumors was significantly increased above that in control animals for two of the fractions (Fractions B and C). Fraction B contained several classes of PAHs (alkylated benzothiophenes, dibenzo- and/or naphthothiophenes, anthracenes and/or phenanthrenes, fluorenes, pyrenes and/or fluoranthenes, benzo- and dibenzofurans, and fluorenones), some of which have documented carcinogenic activity in animal studies (Sivak et al. 1997). Fraction C was reported to contain chemical moieties (e.g. cycloalkenones and/or alkadienones, alkylated phenylethanones, dihydrofuranones, dihydroindenones, isobenzofuranones, hydroxybenzenethiols, pyrenes and/or fluoranthenes, chrysenes and tricarbocyclic fused-ring thiophenes) not previously suspected of being carcinogenic (reviewed in Clark et al. 2011; Sivak et al. 1997). Sivak et al. (1997) speculated that methylated 4- or 5-ring PAHs may have been present in Fraction C, but they were unable to isolate the primary contributors in BURA fumes to the carcinogenic effects observed in mouse skin. Sivak et al. (1997) also conducted a parallel examination of B[*a*]P at several doses in order to compare BURA responses to a known carcinogen (i.e. a positive control).

Sivak et al. (1997) used a laboratory fume generation system that produced fume condensates that differed in composition from those of typical worker exposures (Kriech et al. 2007; NIOSH 2000), In a third chronic skin-painting bioassay, Clark et al. (2011) evaluated the carcinogenic potential of fume condensates generated from a technique more representative of exposures experienced during BURA applications (i.e. field-matched fume), but the bioassay method was otherwise set to be similar to that of Sivak et al. (1997). Specifically, Clark et al. evaluated the carcinogenic response of mice exposed to a field-fume condensate, a laboratory-generated fume condensate, and a parallel examination of B[*a*]P (i.e. positive control). Clark et al. did not, however, evaluate the individual fractions of the asphalt fumes. Thus, the Sivak et al. (1997) and Clark et al. (2011) skin-painting studies each have decided advantages and disadvantages for use in quantitative risk assessment. Examining them together provides a unique opportunity to examine the effects of BURA potency under varying environmental conditions [i.e. varying temperature and varying exposure (field vs. laboratory)] that can influence the quantity and composition of PAHs emitted from BURA (see the "Asphalt chemistry and composition" section).

Quantitative dose–response assessment

In this section, we consider how available animal bioassay data can be used to produce an estimate of BURA fume carcinogenic potency that can then be applied to assess roofer risks. Using the experimental data available, we consider both dermal contact as a potential cause of skin tumors (which is the outcome observed in skin-painting studies on mice) and inhalation as a potential cause of lung tumors (which is the observed outcome in the available inhalation bioassay on rats). Ideally, a toxicity study would include several doses of each tested agent, so that the shape of the dose–response curve could be estimated from several observations. From this dose–response pattern, quantitative models could be fitted to the underlying data and used to extrapolate to a benchmark response (e.g. BMR = 10% tumor incidence). As described in the "Experimental animal data" section, however, available dermal bioassays (Clark et al. 2011; Niemeier et al. 1988; Sivak et al. 1997) employed single doses of each tested asphalt fume condensate, and the studies included both negative and positive (B[*a*]P) controls. Although the available data provide individual animal responses and time of appearance of skin tumors, the single asphalt exposed group for any particular substance precludes comparing response frequency across a range of doses. Estimating potency from a single observation of a high dose, and its high rate of response, risks underestimating the potency at lower doses (at which the saturation of possible response does not occur), but it simultaneously

risks overestimating potency at lower doses (because the nonlinearity of the dose–response leads to less-than-proportional risks at lower doses). Resolving the impact of these biases toward either under- or over-estimation entails examining the whole dose–response curve. Our approach, as described more fully below and in our online Supplementary materials, is to try to obtain information about the shape of the dose–response curve by evaluating the skin-tumor results for the simultaneously tested B[*a*]P positive controls, which included several doses and presumed that the dose–response curves were parallel.

Further, the assessment by Sivak et al. (1997) of skin tumors from various fractions of the whole fume condensate, tested alone and in various combinations, enables examination of tumor rates (and time-to-tumor appearance) for different components of the entire asphalt mixture. Together these data provide a richer evaluation of fume-condensate effects that benefits from the value of partial replication. They can also be used (along with the results from Clark et al. 2011) to illuminate the potential effects of compositional variation. But to realize these benefits, requires that relative potency among the various substances (differing whole-fume mixtures, their fractions and B[*a*]P positive controls) can be measured in a consistent way that does not depend on the particular timepoint (in view of the increase in response over time) or on the particular dose level (in view of the nonlinear dose–response and the saturation of response at high dose levels) at which the relative potency is measured. Such consistency can be had if one assumes that each substance or fraction-combination shares a common dose–time–response relationship, differing only in a consistent substance-specific potency, such that the response relationships are parallel in both the dose and time dimensions.

To compare the potencies of B[*a*]P and the various PAH mixtures as seen in the asphalt-fume condensates, one must assume that PAH compounds exhibit similar modes of action, affect similar toxicological endpoints, and maintain similar potencies across species and exposure routes (EPA 2010; Nisbet & Lagoy 1992). These assumptions have been reviewed (and in some cases challenged) by numerous authors (e.g. EPA 1993, 2010; Muller et al. 1997; Nagao et al. 1993; Nisbet & Lagoy 1992; Parrott et al. 1995; Putzrath 1997; Silkworth et al. 1995; Tallarida 2006). PAH-induced carcinogenicity generally follows a consistent pattern: (a) oxidative metabolism resulting in reactive intermediates that covalently bind with DNA, (b) formation of DNA adducts, (c) tumor initiation due to DNA mutations, and (d) tumor promotion through a variety of events (EPA 2010). Data from *in vitro* and *in vivo* studies provide evidence that a common mode of action exists for several individual PAHs and some PAH mixtures (EPA 1993, 2010; Nisbet & Lagoy 1992). Further, limited evidence suggests that tumor formation is similar among animal species and by route of exposure (EPA 2010; Nisbet & Lagoy 1992). As acknowledged by EPA, however, several uncertainties remain with extrapolation among PAHs. Nevertheless, we have applied this working assumption to our analysis (consistent with EPA current practices; EPA 1993, 2010) and evaluated existing mixture data to determine whether using this approach for heated asphalt mixtures is reasonable.

Dose–response assessment methods

We proceeded through a series of analyses to understand the nature of the tumorigenic response of BURA (and its components) from the available animal bioassays. Further, we tested a series of assumptions regarding the parallelism of the dose–time–response curves for asphalt mixtures and B[*a*]P to extrapolate the dose–response relationship from B[*a*]P to asphalt mixtures (for which only single-dose studies are available). We primarily used EPA's Multistage Weibull (MSW) Time-to-Tumor model to analyze the probability of neoplastic lesions at specified times (Battelle 2010). Details on the MSW model theory, model parameters, and modeling results are presented in the Supplementary appendix. We conducted these analyses in a manner consistent with EPA's recommended dose–response procedures (EPA 2005, 2012). Briefly, the MSW model fits responses both to dose (for which it uses a polynomial form, with non-negative parameters $_i$ for increase with increasing dose) and to time (for which it uses an exponential form of increase with increasing time, with exponent *c*). The MSW model has the form:

$$P(d,t) = 1 \quad \exp\left[\quad _0 \quad _1 d \quad _2 d^2 \quad \quad _k d^k \right] \quad (t \quad t_o)^c \,, \tag{1}$$

where $P(d, t)$ represents the probability of a tumor by age t (in bioassay weeks) for dose d. The parameter t_0 represents the time between when a potentially fatal tumor is initiated and when it becomes observable (i.e. latency).

As discussed in the following sections, we extrapolated dose–response relationships from B[*a*]P to BURA fume condensate, assuming that the relationships are parallel (i.e. the relationship between concentration and response is similar, and only the magnitude of the potency changes). This notion (i.e. PAHs act in a similar manner and can be scaled by potency) is the underlying assumption within EPA's RPF framework

(EPA 1993, 2010). Thus, our dose–response analysis followed generally accepted EPA practices, with some modifications to enable us to analyze disparate data sets. We first describe the B[*a*]P dose–response relationship, as this information supports the results from other methods of analysis.

B[a]P dose–response assessment

Because the several tested doses of B[*a*]P in Sivak et al. (1997) provide the only means to define the shape of dose-dependence, our first task is to describe this dependence. The concurrent positive control group in the Sivak et al. (1997) mouse dermal bioassay included three B[*a*]P dose groups. We therefore conducted a full dose–response assessment for B[*a*]P. Individual animal data, including times of tumor appearance, were coded according to EPA's procedures for MSW modeling (see the Supplementary appendix). Prior to MSW modeling, adjustments to the raw animal data (Arthur D. Little, Inc. 1989; Sivak et al. 1997) were necessary to convert the discontinuously applied doses to a continuous level of exposure. First, we determined the total milligrams applied per mouse by multiplying the administered dose (in μg/50 μL application) for each treatment group by a factor of 2 to account for the number of applications per week for the 104 weeks of exposure. The resulting number was then divided by the number of days in 104 weeks ($n = 728$) to determine the average mass applied per day of the study (mg/d). Final average daily B[*a*]P doses (0.0001%, 0.001% and 0.01%) were 0.0143, 0.143 and 1.43 μg/day, respectively. The potency estimates were predicted based on the time that tumors (carcinomas and papillomas) were observed or time of death in individual animals.

As detailed in the Supplementary appendix, the dose-dependence of mouse skin tumors on B[*a*]P dose was markedly nonlinear, and was dominated by the quadratic term ($_2$) of the dose polynomial. We seek a description of B[*a*]P dose-dependence that can be described by one dose parameter (for comparison to the BURA-fume experiments, which, having only one positive dose, can only fit one such parameter). Hence, we refit the B[*a*]P data using a reduced model that had only ($_2$), fixing the other ($_n$) at zero, and found that the fit of this reduced model was hardly distinguishable from that of the full model. Our primary purpose is to use the B[*a*]P fit for comparison to the concurrently produced results for BURA-fumes, but the B[*a*]P result is also of interest on its own, especially in view of the recent EPA (2014a, b) draft assessment for B[*a*]P based in part on their analyses of these same data. Accordingly, we rendered our B[*a*]P fit into a point of departure (using 10% extra risk at 104 weeks as the benchmark response), resulting in a benchmark dose (BMD) of 0.072 μg/d with a lower limit (BMDL) of 0.051 μg/d (Table 10).

By comparison, EPA's (2014a) draft B[*a*]P assessment reported a similar BMDL for the Sivak et al. (1997) study of 0.06 per μg/day based on the MSW time-to-tumor model. Although there is no standard approach for extrapolating dermal exposures from animals to humans, EPA (2014a) examined a variety of approaches and ultimately applied a body weight scaling approach ($BW^{3/4}$) to account for species differences. This method is used to calculate human equivalent doses (HEDs) using a cross-species extrapolation factor (e.g. [70 kg human/0.035 kg mouse]$^{3/4}$) (EPA 2011). (We discuss the uncertainties related to this extrapolation method in the "Cancer risk estimates for roofing workers" section.) EPA developed a human equivalent dermal slope factor (DSF) of 0.006 per μg/day (based on Sivak et al. 1997). Following this approach, our estimate of a DSF, based on an independent time-to-tumor analysis, resulted in a similar value of 0.0066 per μg/day (Table 10). The remaining analyses build upon this B[*a*]P dose–response assessment to inform the shape of the dose–response curve for asphalt mixtures.

A second mouse skin-painting bioassay, Clark et al. (2011) included a B[*a*]P positive control, albeit at a single very high dose (0.05% B[*a*]P, Table 9). Although this study was intended to follow the methods of Sivak et al. (1997) in other respects, the single dose was about fourfold higher than the top B[*a*]P dose used by Sivak et al. Interestingly, as detailed in the Supplementary materials, we found that MSW fits to the Clark et al. B[*a*]P data produce a lower potency than that derived from Sivak et al. (1997), as long as the single $_2$ dose parameter is fitted. This suggests a diminution of realized potency at very high exposure levels, which could be attributed to toxicity or to saturation of metabolic activation.

Dose–response for asphalt mixtures

In general, the whole-BURA fume condensate data sets (Table 9) are subject to the limitation that the skin painting studies included only a single treatment group for each test substance (be it whole-fume condensate or fractions thereof), and the observed tumor incidence was quite high. The single dose prevents direct observation of the (presumably) nonlinear shape of the dose–response relationship (based on the presumption that the nonlinearity seen for B[*a*]P would also be seen for BURA fume samples if different doses had been tested). Moreover, because the maximum possible response rate is 100%, when most of the response rates are high, they

Table 10. Dermal cancer potency estimates for alternative dose–response assessment methods.

Mixture	Critical study	Average daily dose	Point of departure	Modeling approach	DSF (per μg/day)[a]	Relative potency ratio[b]
B[*a*]P	Sivak et al. (1997)	0.0143, 0.143, 1.43 μg/day	BMD_{10}=0.072 μg/day $BMDL_{10}$=0.051 μg/day	MSW time-to-tumor model	6.6×10^{-3}	–
B[*a*]P	Niemeier et al. (1988)	1.43 μg/day	ED_{50}=0.517 mg	NA	–	–
Asphalt Fume (Type I)	Niemeier et al. (1988)	7143 μg/day (232 or 316 °C)	ED_{50}=2520 mg (232 °C) ED_{50}=1975 mg (316 °C)	NA	–	2.05×10^{-4} 2.62×10^{-4}
Asphalt Fume (Type III)	Niemeier et al. (1988)	7143 μg/day (232 or 316 °C)	ED_{50}=2345 mg (232 °C) ED_{50}=2040 mg (316 °C)	NA	–	2.20×10^{-4} 2.53×10^{-4}
BURA Lab Fume (Type III)	Sivak et al. (1997)	7143 μg/day (316 °C)	BMD_{10}=835 μg/day $BMDL_{10}$=684 μg/day	MSW time-to-tumor model	1.46×10^{-4}	8.6×10^{-5}
BURA Field Fume (TR-A)	Clark et al. (2011)	7143 μg/day (199 °C)	BMD_{10}=3933 μg/day $BMDL_{10}$=2883 μg/day	MSW time-to-tumor model	3.47×10^{-5}	1.8×10^{-5}
BURA Lab Fume (LR-A)	Clark et al. (2011)	7143 μg/day (232 °C)	BMD_{10}=1105.5 μg/day $BMDL_{10}$=945 μg/day	MSW time-to-tumor model	1.06×10^{-4}	6.5×10^{-5}

B[*a*]P, benzo[*a*]pyrene; BMD_{10}, benchmark dose associated with a 10% Effect; $BMDL_{10}$, lower 10% CI on the benchmark concentration; ED_{50}, 50% tumor incidence; BURA, built-up roofing asphalt; DSF, dermal slope factor; MSW, multistage-Weibull; NA, not applicable; RPF, relative potency factor. RPF = (0.1/Asphalt Fume POD) divided by (0.1/B[*a*]P POD). Point of departure (POD)=ED_{50} or BMD_{10} (based on the unadjusted animal results, consistent with EPA (2012), when comparing the relative potency between different assays, we used the benchmark dose associated with a 10% effect (BMD_{10}) (i.e. the central estimate, and not the lower confidence limit).

[a]Human equivalent doses (HEDs) estimated using a cross-species extrapolation factor (70 kg human/0.035 kg mouse)$^{(0.75)}$ (as recommended by EPA 2011,2014a,b).

[b]Relative potency estimated as the ratio of the slopes of the asphalt mixtures to the concurrent dose–response assessment for B[*a*]P:.

fail to measure potency, since lower exposures (had they been examined) might also have been found sufficient to nearly saturate the maximum total possible response. Because the asphalt mixture data lack sufficient information to characterize dose–response, we examined whether the dose–response measured from B[*a*]P could used to estimate the mixture results. Specifically, we modeled the whole-BURA and BURA-fraction data sets from Sivak et al. (1997) using a reduced MSW model in which the only dose parameter estimated is $_2$, fixing the other $_n$ at zero. That is, we assumed a quadratic dose-dependence parallel to that of B[*a*]P but differing in estimated potency.

As we have noted, the Sivak et al. (1997) study, in addition to providing time-to-tumor results for whole-BURA fume condensate, also examined tumorigenicity of various fractions of the whole-fume condensate, separately and in various combinations. These data allow one to examine whether the chemical components of BURA act independently (i.e. in a response-additive manner, with the risks from each independent component adding together when the components are present), or in a dose-additive manner, (i.e. doses of components combine, albeit with distinct potencies, to push response up a common dose–response curve), or even in a synergistic or antagonistic manner (in which the presence of some components enhances or diminishes the responses that other components would display if experienced alone). These distinctions are useful in judging the potential effects of compositional variation among fume sources. This examination of the potential for fractions to interact does not produce an estimate of dermal carcinogenicity as such, but it does permit clearer interpretation of how applicable particular dermal bioassays are to particular samples of fume condensate. This clarity enables one to extrapolate results from animal bioassays that used relatively well-characterized BURA samples to other BURA fume sources.

Accordingly, we performed MSW time-to-tumor modeling on each individual asphalt fraction (i.e. components A through E and combinations thereof) and compared the relative potency to the whole-Type III BURA fumes (Arthur D. Little, Inc. 1989; Sivak et al. 1997). The first round of MSW fits for each data set estimated two parameters: the single parameter that describes quadratic dose-dependence ($_2$) and the parameter *c* that, as the exponent of time, describes how quickly the risk of a tumor will increase exponentially over time. As further discussed in the Supplementary appendix, we set all exposures to a dose of one "unit" (i.e. for each BURA fraction, a dose of one unit indicates an amount equal to that fraction's

representation in whole-fume condensate). This setting allows each fraction to be treated as one component of the whole asphalt mixture, which can then be used to test how the components act when the units of each fraction are added together in an exposure. The results of this analysis (as shown in the Supplementary appendix) illustrate that the time-dependence of tumor response (defined by parameter c) is roughly similar among BURA fractions (ranging from 5.8 to 9.5, with a mean of 7.064). (The c parameter from the B[a]P dose–response assessment was also within this range: $c = 6.5$). The uncertainty bounds on each data set's fitted value of c were broad and widely overlapping. This overlap indicates that the value of c is not strongly constrained by the data, and it is consistent with there being a single common value of c across the data sets, with variation from one data set to the next being attributable to sampling variation. As discussed in the Supplementary appendix, a common value of c is valuable, because it renders the dose–time–response patterns parallel not only in the dose dimension (using the assumption of quadratic dose-dependence, as described above) but also in the time dimension. Thus, a relative potency among the various tested substances can be defined that will apply to any chosen dose level or timepoint.

Thus, we adopt the assumption of a common power of time, and we use the observed average among the MSW fits to the various BURA fractions ($c = 7.064$) as its value. We refit the MSW to all the Sivak et al. (1997) BURA-fraction tumor data from mouse skin-painting, fitting the single $_2$ parameter independently for each data set but fixing c at the common value. We also re-analyzed the B[a]P dose–response data using the same common value of c. The results (as shown in the Supplementary appendix) demonstrate that fixing c at the common average value does not appreciably alter the shape of the fitted model or deteriorate its fit to any of the data sets, including the case of the B[a]P positive controls. Of course, the estimated data set-specific values of $_2$ do change once c is fixed, because they are in a sense scaled to the magnitude of c as it appears elsewhere in the equation. (For instance, for B[a]P the original modeled $_2$=1.77 10 ^{12} becomes $_2$=1.56 10 ^{13} after c is fixed.)

In short, we used a simplified equation for a reduced MSW model:

$$P(d,t) = 1 \quad \exp[(\ _2 d^2)\ \ (t \quad t_o)^c\ . \qquad (2)$$

In applying this model, only $_2$ is fitted. ($_0$ is set to zero, since there are no control tumors; t_0 is set to 1, as discussed in the Supplementary appendix; and c is set to the averaged value, as discussed above.) The resulting values of $_2$ are presented in Table 11.

Next, using these results, we compared the estimate of fume fraction potency ($_2$ term) across groups and to the whole BURA fume, and to B[a]P. Three of the five fractions of the BURA fume (fractions A, D and E) showed no carcinogenic effect when mice were exposed to them alone (and hence $_2 = 0$); the basis for whole-fume effects appears to lie with fractions B and C, both of which produced substantial effect when tested alone, albeit an effect smaller than that seen for whole BURA-fume condensate. Notably, the 2- to 3-ringed PAHs, which make up a large fraction of BURA, appear mostly in Fraction A (Sivak et al. 1997). When selected fractions are combined and tested, those containing either B or C, or both, show effects. Although not shown in Table 11, combinations of fractions A, D, and E showed no effects, with the exception that two papillomas appeared for the test of A+D+E combined.

The next question, then, is how the effects of the fractions combine – whether they act independently or additively, and whether there is evidence of interference or enhancement of one fraction's effect when experienced together with another. When using parameter estimates from the reduced model of Equation (2), the relative potency of two substances is given by the square root of the ratio of their $_2$ values. (The square root is used because the $_2$ apply to the square of the dose.) Because the doses are expressed as one "unit" for each tested substance, the relative potencies represent the relative effect of the different components of whole fume condensate when experienced as the amounts that appear in the whole fume condensate itself. Table 12 shows the relative potencies, calculated as square roots of the ratios of the $_2$ values in Table 11, for fraction B alone versus C alone, for fraction B versus each of the exposures that included both B and C (and, depending on the case, also A, D, and/or E as well), and for fraction C versus each of the exposures that included both B and C. It also shows relative potencies of fraction B versus those materials including B but not C, and for fraction C versus those materials including C but not B.

Table 12 shows that fraction B when tested alone is about half as potent as C when tested alone (the estimated relative potency being 0.48). If this is true, and if the presence of any of fractions A, D and E do not alter the relative potency of B and C, then the relative potency of B alone to other tested substances that include both B and C should be about one-third (since one unit of B is being compared to mixtures including one unit of B and one unit of twice-as-potent C, or in effect, three units of B). Table 11 shows some variation among the comparisons, but the average relative potency over the seven possible comparisons is indeed as expected (0.32, see Table 12). By the same reasoning,

Table 11. Comparison of various approaches to estimate the potency of BURA fractions painted on mice from Sivak et al. (1997).

Treatment description[a]	Carcinoma (n)[b]	Time to carcinoma (weeks)	Response (*R*)	MSW Model[c]	
				Potency ($_2$)	Fraction[d] relative potency
Whole Asphalt Fume	20	73.9	0.67	4.66E 14	1.0
0.01% B[*a*]P	27	55.9	0.90	4.17E 13	3.0
Fraction A	0	0	0.00	0	–
Fraction B	10	98.3	0.33	4.08E 15	0.3
Fraction C	17	85.6	0.57	1.76E 14	0.6
Fraction D	0	0	0.00	0	–
Fraction E	0	0	0.00	0	–
Fractions A + B+C + D+E	19	75.4	0.63	6.86E 14	1.2
Fractions A + B	8	97.4	0.27	7.73E 15	0.4
Fractions A + C	11	89.6	0.37	1.58E 14	0.6
Fractions B + C+D + E	15	80.7	0.50	2.41E 14	0.7
Fractions A + B+C + D	18	80	0.60	4.35E 14	1.0
Fractions A + B+C + E	21	77.4	0.70	5.20E 14	1.1
Fractions B + C+D	15	86.4	0.50	2.62E 14	0.8
Fractions B + C	24	73.4	0.80	5.15E 14	1.1
Fractions A + C+D + E	14	88.6	0.47	1.22E 14	0.5
Fractions A + B+D + E	7	99.3	0.23	3.13E 15	0.3

B[*a*]P, benzo[a]pyrene; BURA, built-up roofing asphalt; MSW, multistage-Weibull; *N*, number. All methods assume a dose of 1 unit, or that a complete dose from each fraction, from a given amount of asphalt fume, is administered to each animal.
[a]Fractions A–E represent different chemical groups fractionated from a raw BURA Type III asphalt fume.
[b]Defined as an animal with at least one carcinoma (Arthur D. Little, Inc. 1989).
[c]Potency estimated using the $_2$ term predicted from the MSW time-to-tumor model based on the time of observation of tumors (carcinomas and papillomas), or time of death in individual animals. The following parameters were specified in the MSW model: c=7.064, $t\_0 = 1$, beta_0 = 0, beta_1 = 0 (see the Supplementary appendix for details on model parameters).
[d]Fraction relative potency estimated as the square root of ($_2$ of each fraction/ $_2$ of whole asphalt fume).

Table 12. Relative potency of asphalt fume fractions.

Description[a]	Relative potency[b]
B to C	0.48
B to whole	0.30
B to A + B+C + D+E	0.24
B to B + C+D + E	0.41
B to A + B+C + D	0.31
B to A + B+C + E	0.28
B to B + C+D	0.39
B to B + C	0.28
Average	0.32
C to whole	0.61
C to A + B+C + D+E	0.51
C to B + C+D + E	0.86
C to A + B+C + D	0.64
C to A + B+C + E	0.58
C to B + C+D	0.82
C to B + C	0.58
Average	0.66
B to A + B+D + E	1.14
B to A + B	0.73
Average	0.93
C to A + C+D + E	1.20
C to A + C	1.06
Average	1.13

[a]Fractions A–E represent different chemical groups fractionated from a raw BURA Type III asphalt fume from Sivak et al. (1997).
[b]Relative potency calculated as the square root of the ratio of the potency of respective fractions: Relative Potency = $\sqrt{(}$ $_{2\text{-Fraction B}}$/ $_{2\text{-Fraction C}}$). $_2$ values for each fraction provided in the Supplemental appendix and Table 11.

comparisons of C alone to mixtures including both B and C should show a relative potency of about two-thirds (one unit of C being compared to one unit of half-as-potent B plus one unit of C, or in effect, 1.5 units of C), and on average over the seven relevant comparisons, the value obtained is 0.66 (Table 12). Similarly, if A, D and E do not interfere or enhance effects, B alone should be equally potent to mixtures including B but not C (and the observed average of the two applicable observations is 0.93) and C alone should be equally potent to mixtures including C but not B (and the observed average is 1.13).

This is the pattern expected under dose-additivity, in which the fractions B and C can be treated as dilutions of one another according to their constant relative potency, and an exposure to both fractions has a response expected from the appropriately scaled sum of the component exposures. Individual comparisons vary somewhat, which is to be expected owing to sampling error, but the overall pattern is robust. It appears that fractions B and C add together in producing their effect, that they do not interact synergistically or antagonistically, and that the additional presence of fractions A, D and E neither enhances nor diminishes the effects from the B and C fractions.

Another way to investigate this question is to examine the risk outcomes. The caveat is that the raw end-of-study response rates (number of tumor-bearing animals as a fraction of those tested) is not corrected for intercurrent mortality, and some mice that eventually might have developed tumors may have been removed before tumors had a chance to appear. Also, because response rates are high, the possibility that response was saturated should be considered. Our approach was to use the fitted MSW model equations for each tested mixture to predict responses, $P(d, t)$, at two fixed times

Table 13. Modeled risk predictions for asphalt fume fractions.

	70 Weeks		90 Weeks	
Treatment description	Model risk prediction[a]	From components[b]	Model risk prediction[a]	From components[b]
Whole asphalt fume	0.366		0.936	
0.01% B[*a*]P	0.983		1.000	
Fraction A	0.000	0.000	0.000	0.000
Fraction B	0.039	0.039	0.214	0.214
Fraction C	0.158	0.158	0.646	0.646
Fraction D	0.000	0.000	0.000	0.000
Fraction E	0.000	0.000	0.000	0.000
Fractions A + B+C + D+E[c]	0.488	0.191	0.983	0.721
Fractions A + B	0.073	0.039	0.366	0.214
Fractions A + C	0.143	0.158	0.606	0.646
Fractions B + C+D + E[c]	0.209	0.191	0.758	0.721
Fractions A + B+C + D[c]	0.346	0.191	0.923	0.721
Fractions A + B+C + E[c]	0.398	0.191	0.953	0.721
Fractions B + C+D[c]	0.226	0.191	0.787	0.721
Fractions B + C[c]	0.395	0.191	0.952	0.721
Fractions A + C+D + E	0.113	0.158	0.514	0.646
Fractions A + B+D + E	0.030	0.039	0.168	0.214
Average for Fractions with B + C[c]	0.344		0.893	

Fractions A–E represent different chemical groups fractionated from a raw BURA Type III asphalt fume from Sivak et al. (1997).

[a]Risk predictions estimated using Equation (2) $=P(d,t)=1 \quad \exp[(\ _2d^2) * (t \quad t_0)^c])$, where $_2$, the potency for each fraction provided in the Supplemental appendix and Table 11; d, dose set at 1 unit of exposure; *t*, time in weeks (70 or 90 weeks); t_0, latency time from tumor initiation to observation (1 week); *c*, the exponent of time (average value of 7.064, as discussed in the Supplemental appendix).

[b]Represents the risks of Fraction B and Fraction C when examined alone, combined according to independent action (response additivity).

[c]Represents the average of risk predictions for mixture samples including both fractions B + C.

(70 weeks and 90 weeks) corrected for intercurrent mortality. These predictions are tabulated in Table 13.

At 70 weeks, the average of all mixture-specific MSW equation risk predictions for mixtures including both B and C is 0.344, and at 90 weeks it is 0.893. If one uses the relative potency of B–C being 0.48, as already calculated, then an exposure to B and C together results, in "B-equivalents", in an exposure of $1 + (1/0.48) = 3.08$ units of B. If this result is put into the MSW equation fitted to the test of B alone, it yields a 70-week risk of 0.31 and a 90-week risk of 0.90, in good agreement with the observed average over all the MSW model calculations for B-and-C-containing mixtures. Similarly, an exposure to B and C in "C-equivalents" is $0.48 + 1 = 1.48$, and calculating the MSW equation outcomes from the C-only fit for a dose of 1.48 yields a 70-week risk of 0.31 and a 90-week risk of 0.90 – again in good agreement with the averages for models fit to the mixtures themselves. This constitutes a direct evaluation of the success of dose-additivity among the data available to test it.

One can contrast this dose-additivity result with that predicted by response addition. If the effects of B and C are not additive, but each has an independent effect, then the risk from experiencing both should be the sum of the risks of each alone (minus their product, to account for the overlap in some animals of an effect from B and an independent effect from C). (If the relation between dose and response is strictly linear, dose-addition and response-addition are equivalent, but if dose–response is nonlinear, as in this case, the results differ). From Table 12, one can take the responses for B alone and for C alone and calculate the expectation under independence (i.e. response addition). These are $0.04 + 0.16 \quad (0.04 \quad 0.16) = 0.19$ for 70 weeks and $0.21 + 0.65 \quad (0.21 \quad 0.25) = 0.72$ for 90 weeks. These results modestly but systematically underestimate the observed 70- and 90-week responses from mixture-specific model equations of 0.32 and 0.88, respectively. That is, is the response was somewhat greater when B and C were experienced together by the Sivak et al. (1997) bioassay mice than would be expected from combined independent responses to each of the components, but the results are very much in accord with the expectation of dose-addition. The higher response for B + C from dose-addition than from response-addition reflects the nonlinearity of the dose-dependence, with dose-addition increasing the response more than proportionally as the combined dose pushes the response up the nonlinear curve. We conclude that the hypothesis that the components of the BURA-fume-condensate as tested by Sivak et al. (1997) are dose-additive, and that the presence or absence of fractions that are not in themselves carcinogenic neither enhanced nor suppressed this response, is supported.

Dermal carcinogenicity potency estimation

Having analyzed the bioassay results for skin-painting of BURA-fume condensate in mice, we considered how to apply these results to defining a basis for estimating the

possible magnitude of skin cancer risk in humans. We considered three general quantitative approaches to defining potency of BURA fumes for skin cancer:

(1) *Whole-mixture assessment*: We used the bioassay results as a direct observation on dermal carcinogenicity for the whole BURA Type III mixture. This approach entails estimating the potency in mice, defining an appropriate point of departure for low-exposure risk characterization, and extrapolating this effect to humans by considering dermal exposures that should be presumed to be of equivalent risk in the two species. This method implements the first (and preferred, when feasible) choice of EPA's mixtures guidelines (EPA 2000), to base assessment of a mixture on testing the mixture itself.
(2) *Potency relative to concurrent B[a]P groups*: We evaluated the potency of whole asphalt fumes relative to the potency of B[*a*]P when tested in parallel – that is, in the same mouse model, with equivalent observations of the same tumors. Doing so in effect defines the asphalt fume exposures in terms of B[*a*]P-equivalents. An externally supplied assessment of the human dermal carcinogenicity of B[*a*]P can then be used to estimate human risks based on the B[*a*]P-equivalents experienced in an exposure to BURA fumes. Clearly, in addition to the observation of relative potency, this approach hinges on the validity of the chosen basis for projecting human dermal cancer risks from B[*a*]P.
(3) *Relative potency factors (RPF approach)*: We applied the existing RPF methods for specific PAH components to the concentrations of those PAHs in BURA fume condensate. Existing RPF methods define potencies relative to B[*a*]P as an index chemical, with each specific RPF assigned based on expert judgment in view of the existing database of information about that PAH's toxicity compared to that of B[*a*]P, such that each component contributes an effect to the whole gauged in B[*a*]P-equivalents. Following the usual application of the RPF methods, we assumed that the overall effect of the mixture is attributable to the sum of the independent actions of those components that have defined RPFs, that the RPFs that have been developed accurately estimate such independent effects, and that those polycyclic aromatic compounds without defined RPFs do not contribute to, enhance, or suppress the effects of the PAHs that have defined RPFs.

We explored these approaches to quantify dermal and inhalation cancer potency for heated asphalt mixtures.

Whole-mixture assessment

Three dermal carcinogenicity bioassays are available for heated asphalt mixtures (Clark et al. 2011; Niemeier et al. 1988; Sivak et al. 1997). Niemeier et al. (1988) examined one dose each of various asphalt and coal tar pitch condensates, as well as B[*a*]P as a positive control. These authors roughly characterized relative potency by reporting the cumulative amounts of total exposure (summed progressively over weeks) up to the time point at which 50% tumor incidence (ED_{50}) was reached in the mice for each group (Table 10). The design of the Niemeier et al. (1988) study does not lend itself to a direct measurement of potency, and hence these data are considered under the "potency-relative-to-concurrent- B[*a*]P" approach discussed below. In contrast to Niemeier et al. (1988), the carcinogenicity bioassays performed by Clark et al. (2011) and Sivak et al. (1997) included the individual animal data. Therefore, a detailed analysis of the whole asphalt mixtures was practicable using the MSW modeling procedure discussed in the "Dose–response assessment methods" section. Specifically, we examined whole BURA fume mixtures, including the lab generated fume (Type III) from Sivak et al. (1997) and the field (TR-A) and lab (LR-A) generated fumes from Clark et al. Because these studies included only one dose group for each mixture, we imposed the fixed-dose–response parameters ($c = 7.064$, $t = 104$ weeks, and $t_0 = 1$) and estimated a mixture-specific $_2$ (as discussed in the "Dose–response assessment methods" section and the Supplementary appendix). The MSW model was fitted to the TWA dermal dose rates from the Sivak et al. (1997) and Clark et al. (2011) studies using the doses and individual animal data (day of death, or first day of carcinoma or papilloma observation).

The laboratory-generated fume (reported by Sivak et al. 1997) was generated under conditions similar to those used by Niemeier et al. (1988), and the Clark et al. (2011) study aimed to replicate the fume-generation and testing methods of the earlier studies. Notably, these two near-replicate studies (Clark et al. 2011; Sivak et al. 1997) produced very similar estimates of potency based on fitted MSW models. As noted in the "Occupational inhalation exposure" section, however, the similarity of these laboratory-generated fumes to those experienced by workers in the field has been questioned. Clark et al. also examined field-generated asphalt fume (TR-A) and found that it was less potent than the laboratory-generated fumes (LR-A or Type III BURA). The field-generated fume (TR-A) is approximately 78% less potent than the laboratory fume generated by Clark et al. (2011) and 24% less potent than the laboratory fume generated

by Sivak et al. (1997). These results are consistent with mutagenicity assays comparing field- versus laboratory-generated fumes (using similar fume generation methods) that indicated that field-generated fumes had a mutagenicity index (MI) of 0.0 compared to laboratory-generated fumes with MIs of 5.3 (generated at 149 °C) and 8.3 (generated at 316 °C) (Reinke et al. 2000). This comparison indicates that compositional (and potency) differences among BURA fume samples is, in part, a function of the fume generation temperature.

Having characterized the dose–response relationships and potencies as observed in the mouse skin-painting studies, it remained to express these results in a form that can be applied to estimating human dermal-exposure risk. This task has two components: (a) defining a point of departure for inferences about low exposures, and (b) evaluating toxicologically equivalent exposure measures for mice in the experiments and humans exposed during roofing work.

For the first of these components, following EPA methods for defining potency applicable to low exposure levels, we used a linear low-exposure extrapolation from the lower 10% confidence level on the benchmark concentration ($BMDL_{10}$) when deriving potency estimates (EPA 2005). This is an extrapolation downward to low doses that are not directly observable in the bioassay, but starting the linear extrapolation from a low response level preserves the information about non-linearity of dose–response at higher doses. This approach reflects the recognition that extrapolation below the observable range is largely a policy choice, and that a linear extrapolation defined in this way is expected to be a conservative (health protective) assumption in the face of that uncertainty.

The second component – determining a human exposure of equivalent effect to the exposures tested in mice – is particularly challenging for dermal exposures, for which established methods and policies for applying them have not been well defined. This component in turn has two aspects. The first is the issue of scaling (to account for mouse:human differences in mass, skin surface, the pace of relevant physiological processes, and the difference in lifespan). Existing practices for cross-species dose scaling are based on rationales developed for internal doses. In its draft assessment of B[*a*]P, EPA (2014a, b) considers this scaling question at some length and entertains a series of possible options (e.g. risk proportional to mass, surface area, or body weight). After reviewing these alternatives, consistent with EPA's determination for B[*a*]P (EPA 2014a, b), we have adopted the allometric scaling ($BW^{3/4}$) cross-species extrapolation procedure. Extrapolation was accomplished by multiplying the POD by the $(BW_{human}/BW_{Mouse})^{3/4}$, where the mouse and human body weights were EPA default reference body weights $(70\,kg/0.035\,kg)^{3/4}$ (Table 10).

The other aspect of human equivalence is the relative susceptibility of human versus mouse skin to carcinogenic effects of PAHs (Atillasoy et al. 1997, 1998; Balmain & Harris 2000; Berking et al. 2002; Graem 1986; Kappes et al. 2004; Roelofzen et al. 2010; Soballe et al. 1996; Urano et al. 1995). In following the EPA-suggested method, we considered skin to be equivalently susceptible across species (as long as exposures are corrected for scale, as noted above), but there are in fact indications that human skin is systematically less susceptible even to scale-adjusted exposures. This question is further discussed in the "Cancer risk estimates for roofing workers" section as we apply the evaluations to estimates of roofer dermal exposure. Briefly, though considerations of metabolic activity suggest notably lower activation in human than in mouse skin; xenografts of human skin onto mouse backs show much less response than mouse skin in skin-painting studies of other PAH mixtures, and epidemiological evidence from historical use of coal tar as an eczema treatment does not show skin cancer risks that would be expected if mouse-human equivalence in susceptibility were maintained (Atillasoy et al. 1997; Balmain & Harris 2000; Pittelkow et al. 1981; Soballe et al. 1996; Urano et al. 1995). Thus, we estimated the potencies and DSFs (Table 10) for three similar whole asphalt mixtures to be applied in evaluating dermal cancer risks to roofing workers in the "Cancer risk estimates for roofing workers" section.

Potency relative to concurrent B[a]P groups

We compared potency estimates from fractionated BURA samples (Table 10) to potency estimates from concurrent B[*a*]P treatments from the same experiments. In the Niemeier et al. (1988) study, based on comparison of the total solids concentration, tumorigenic response in the control group, which was exposed to B[*a*]P, was over 1000 times greater than in animals exposed to asphalts (Type I or III) and over 100 times greater than those exposed to coal tars (as measured by the relative total accumulated mg of the mixtures to total accumulated mg of B[*a*]P needed to produce 50% response). The lab-generated fumes collected at higher temperatures (316 °C) were generally more potent carcinogens than those collected at lower temperatures (as indicated by the lower ED_{50} values at higher temperatures; Table 10). Niemeier et al. (1988) suggested that B[*a*]P content was not an adequate measure to characterize

the tumorigenic response of asphalts, possibly owing to the presence of other aliphatic hydrocarbons.

The bioassays by Sivak et al. (1997) and Clark et al. (2011), by supporting MSW model fitting, permit a more quantitative assessment of PAH potency relative to that of B[*a*]P. In general, the MSW-estimated potencies are consistent with results from Niemeier et al. (1988), in that asphalt fumes (field- or laboratory-generated) were less potent than B[*a*]P. As shown in Table 10, B[*a*]P (from concurrent study groups) was orders of magnitude more potent than asphalt fumes (from Niemeier et al. 1988; Sivak et al. 1997; or Clark et al. 2011). This potency difference is likely related to the chemical composition of the mixtures and also the relative rates of absorption as compared to the B[*a*]P controls (see discussion in the "Occupational dermal exposure" section).

RPF approach

The most commonly used application of an RPF-based approach involves evaluating mixtures of PAHs based on the relative potency of individual PAHs to B[*a*]P and considering the risk expected from the resulting combined exposure in terms of B[*a*]P-equivalents. Given the prominence of this method, coupled with the available data on the levels of PAHs in BURA fumes, we applied the PAH RPF approach to the PAH content of asphalt fumes (Table 14), using concentrations of 16 PAHs that were described for the BURA fume condensates used in the skin-painting studies (Kriech et al. 2007; Sivak et al. 1997) and applying three RPF approaches (EPA 1993, 2010; WHO/IPCS 1998) to assess the potency. When using this method, we expressed doses of each individual constituent in a mixture that has a defined RPF in terms of B[*a*]P-equivalents, added these individual doses together, and estimated the response for the total B[*a*]P-equivalent dose using the method's supplied approach to estimating risk for B[*a*]P dermal exposures in humans (EPA 1993, 2010; WHO/IPCS 1998). EPA (2010) recommends these types of component-based approaches when toxicity data on complex mixtures or "sufficiently similar" mixtures are not available.

The RPF analysis (Table 14) demonstrated that PAH concentrations in the laboratory-generated whole BURA fume are very low, resulting in the poor predictive ability of the PAH-RPF method in estimating the tumor response in mice (i.e. 0.9 expected vs. 20 observed, 3.8 expected vs. 35 observed). The RPF prediction for the field-generated fume (TR-A) provided a closer estimate compared to the observed tumor response (i.e. 2.0 expected vs. 8 observed). We note that these results are based on EPA's 1993 set of RPFs (Table 14). Other sets of RPFs exist in either draft or final form (e.g. EPA 2010; WHO/IPCS 1998). We tested these alternative RPFs (data tables not shown here) and found the results to be similar to those of EPA's 1993 RPF set. Specifically, using EPA's Draft 2010 RPFs the expected versus. observed tumor response was 3.7 versus 20 (Lab Type III), 4.2 versus 8 (Field TR-A), and 8.1 versus 35 (Lab LR-A). Similarly, using WHO's RPFs the expected versus observed tumor response was 1.1 versus 20 (Lab Type III), 2.5 versus 8 (Field TR-A) and 4.5 versus 35 (Lab LR-A). As previous investigators have observed, these data demonstrate that PAH content is not an effective determinant of the carcinogenic activity of BURA fume (Niemeier et al. 1988; Sivak et al. 1997), and, in view of the many and diverse polynuclear aromatic compounds in BURA fume, the total effect likely includes contributions from a number of constituents that do not have established RPFs and are therefore not counted (Clark et al. 2011).

Assessment of fluorescence and dermal carcinogenicity

Some degree of variation has been demonstrated in composition of the roofing asphalt mixtures, and the composition of fumes can vary with such differences as well as with the method and temperature of fume generation. We have also found that these variations appear to make modest alterations in skin-cancer potency. Performing long-term animal studies on every mixture variant of interest would be impractical, so it is useful to explore rapid assessment methods for adjusting dermal carcinogenic potency estimates that may result from compositional variation. One such method is the NIOSH Method 5800 (NIOSH 1998), which uses an ultraviolet (UV) fluorescence technique to examine the fluorescence of illuminated samples in a way tuned to detect the amounts of the multi-ringed PAH fractions that are primarily responsible for carcinogenic effects. Osborn et al. (2001) developed a modified fluorescence technique to maximize the signal from compounds present in the fume fractions that increased skin tumor incidence in the Sivak et al. (1997) study. Several authors have studied the relationship between Osborn UV fluorescence [in emission units (EU) per gram of material] and carcinogenicity using the results from the Sivak et al. (1997) bioassay (Osborn et al. 2001; Kriech et al. 1999; Trumbore et al. 2011). Carcinogenicity was quantified using an index (CI/g = [Carcinoma Incidence/time in weeks to first carcinoma 100]/g of BURA fraction). According to Kriech et al. (1999), the index is arbitrarily multiplied by 10 to give convenient values. The carcinogenicity index (CI/g) has an advantage over simple incidence-based metrics in that it includes a temporal

Table 14. Comparison of the predicted carcinogenicity of BURA fumes in mice using established RPFs versus the observed tumor incidence in the animal bioassay.

		Lab-BURA (Type III)[a]		Field-BURA (TR-A)[a]		Lab-BURA (LR-A)[a]	
PAH	EPA RPF[b]	PAH content (mg/kg)	B[*a*]P$_{eq}$ risk[c]	PAH content (mg/kg)	B[*a*]P$_{eq}$ risk[c]	PAH content (mg/kg)	B[*a*]P$_{eq}$ risk[c]
Acenaphthene	–	NA	–	NA	–	NA	–
Acenaphthylene	–	NA	–	16.0	–	18.0	–
Anthracene[d]	–	NA	–	23.0	–	33.0	–
Benz[*a*]anthracene	0.1	8.0	5.56E 03	12.0	8.33E 03	20.0	1.39E 02
Benzo[*a*]pyrene	1.0	3.0	2.08E 02	3.1	2.15E 02	5.8	4.03E 02
Benzo[*b*]fluoranthene	0.1	5.0	3.47E 03	2.8	1.94E 03	5.5	3.82E 03
Benzo[*e*]pyrene	–	4.0	–	8.1	–	17.0	–
Benzo[*g,h,i*]perylene	–	1.0	–	2.0	–	2.8	–
Benzo[*k*]fluoranthene	0.01	NA	–	1.9	1.32E 04	2.9	2.01E 04
Chrysene	0.001	13.0	9.03E 05	14.0	9.72E 05	20.0	1.39E 04
Dibenz[*a,h*]anthracene	1.0	NA	–	NA	–	NA	–
Fluoranthene	–	97.0	–	6.8	–	22.0	–
Fluorene	–	39.0	–	110.0	–	82.0	–
Indeno[1,2,3-*cd*]pyrene	0.1	2.0	1.39E 03	0.9	6.04E 04	1.2	8.33E 04
Naphthalene	–	17.0	–	110.0	–	39.0	–
Phenanthrene	–	300.0	–	130.0	–	250.0	–
Pyrene	–	63.0	–	25.0	–	150.0	–
Total risk of tumor in mice			0.031		0.033		0.059
Observed tumor incidence		67%		13%		55%	
Observed number of tumors[e]		20 ($n = 30$)		8 ($n = 62$)		35 ($n = 64$)	
Expected number of tumors[f]		0.9		2.0		3.8	

B[*a*]P, benzo[*a*]pyrene; B[*a*]P$_{eq}$, benzo[*a*]pyrene equivalent; BURA, built-up roofing asphalt; EPA, United States Environmental Protection Agency; LR, lab roofing; *n*, sample size; NA, not available; PAH, polycyclic aromatic hydrocarbon; RPF, relative potency factor; TR, tank roofing.

[a]Mean PAH concentration (ppm) as reported by Sivak et al. (1997) (lab-generated BURA Type III sample, whole asphalt fume fraction) and Kriech et al. (2007) (field and lab-generated BURA samples).

[b]EPA (1993) RPFs.

[c]B[*a*]P$_{eq}$ risk estimated as PAH content EPA RPF average daily dose (ADD) of 0.005 μg/day/mouse B[*a*]P DSF (0.1/0.072 = 1.39 per μg/day for mice from Table 10).

[d]Concentration reported as anthracene/phenanthrene in Arthur D. Little, Inc. (1989).

[e]Tumor incidence for Lab-BURA (Type III) from Sivak et al. (1997) and for TR-A and LR-A from Clark et al. (2011).

[f]Expected number of tumors in mice = total risk Number of animals.

component (i.e. time in weeks to first tumor). This approach ascribes a higher potency to a group response with a shorter latency (i.e. more aggressive tumors) compared to an index that considers only the incidence of animals with tumors relative to animals at risk. Fluorescence provides a measure of the toxicity of the whole BURA fume itself, rather than assessing the toxicity of its individual components. Osborn et al. (2001) proposed a fluorescence technique as a screening tool for assessing biologically active PAHs in asphalt fumes, based on the following observations:

> The response was highly correlated with the presence of 4- to 6-ring PAHs that were identified as the primary chemical fractions associated with tumor formation in the asphalt bioassays (Niemeier et al. 1988; Sivak et al. 1997).
>
> Examination of individual fractions versus the summation of fractions demonstrated that the fluorescence method was reliable (e.g. individual EU/g results were 81, 92, 167, 7.2 and 0.5 for Fractions A, B, C, D and E, respectively, and the total of 348 EU/g compares to the whole fume result of 325 with 107% recovery).
>
> Smaller ringed PAHs (2–3 rings) influenced the fluorescence response, and correction for these PAHs reduced the fluorescence signal (e.g. Fraction A reduced from 124 to 39 EU/g) but did not diminish the response of 4- to 6-ring PAHs.
>
> Testing of other oil products found that very viscous oils (with 7-ring PAHs) and elevated distillation temperatures were not well described by this relationship. Therefore, the carcinogenic response predicted by this method is not reliable for asphalts or other oil-derived products outside the distillation range of paving and roofing fumes (187–346 °C).

Similar relationships were described by both Kriech et al. (1999) and Trumbore et al. (2011) using alternative fluorescence measurements for the Sivak et al. (1997) BURA fractions. Trumbore et al. (2011) averaged fluorescence measures for the Sivak et al. (1997) fractions, where several measurements were available. In all cases, a strong correlation was obtained between the two measures, regardless of the source (or combination) of fluorescence measurements.

We examined the EU versus CI correlation using the Osborn et al. data set (with the average EUs for each

Table 15. Comparison of measured and predicted potency by UV fluorescence.

Fume fraction[a]	EU/g[b]	MSW potency ($_2$)[c]	Grams painted[d]	Normalized potency[e]	Predicted potency[f]	Ratio of predicted potency relative to whole fume
Whole asphalt fume (Type III)	326	4.66E 14	5.2	4.15E 08	5.78E 08	1
Fraction A	156	0	2.79	0	6.80E 09	0.1
Fraction B	1856	4.08E 15	0.13	4.91E 07	5.17E 07	8.9
Fraction C	1788	1.76E 14	0.25	5.31E 07	4.96E 07	8.6
Fraction D	42	0	0.44	0	NA[g]	NA
Fraction E	11	0	0.13	0	NA[g]	NA
Fraction A + B	256	7.73E 15	2.79	3.15E 08	3.68E 08	0.6
Fraction A + B+C + D	308	4.35E 14	3.57	5.84E 08	5.24E 08	0.9
Fractions A + B+C + D+E	299	6.86E 14	3.67	7.14E 08	4.97E 08	0.9
Fractions A + B+C + E	330	5.20E 14	3.32	6.87E 08	5.90E 08	1
Fractions A + B+D + E	205	3.13E 15	3.33	1.68E 08	2.15E 08	0.4
Fractions A + C	293	1.58E 14	3.03	4.15E 08	4.79E 08	0.8
Fractions A + C+D + E	308	1.22E 14	3.44	3.22E 08	5.24E 08	0.9
Fractions B + C	1395	5.15E 14	0.62	3.66E 07	3.79E 07	6.5
Fractions B + C+D	718	2.62E 14	0.82	1.98E 07	1.75E 07	3
Fractions B + C+D + E	619	2.41E 14	1.29	1.20E 07	1.46E 07	2.5
Field roofing fume (TR-A)	157	5.65E 17	5.2	1.45E 09	7.10E 09	0.1
Lab roofing fume (LR-A)	336	5.54E 15	5.2	1.43E 08	6.08E 08	1.1

BURA, built-up roofing asphalt; EU, emission unit; MSW, multistage-Weibull; NA, not applicable; UV, ultraviolet.
[a]Fractions A–E represent different chemical groups fractionated from a raw BURA Type III asphalt fume.
[b]UV fluorescence measurements adapted from (Kriech et al. 1999, 2007; Osborn et al. 2001; Trumbore et al. 2011).
[c]Potency estimates as described in Table 15.
[d]Dosing amounts adapted from Sivak et al. (1997) and Clark et al. (2011).
[e]Normalized potency calculated according to the square root of (B_2/mass^2).
[f]Predicted by relationship presented in Figure 4 [y=3 10 ^{10}(x)–4 10 ^{8}].
[g]NA = not applicable; the fluorescence measure falls below the range of the regression function (below 130 EU/g).

fraction) and compared these to an alternative in which we compared EU to our own determination of potencies (using the MSW time-to-tumor analyses described above) for the Sivak et al. (1997) asphalt fume fractions (Table 15). Specifically, we examined the potency of the fraction (described by the $_2$ term) normalized to the amount of grams painted in the mouse dermal assays, using the following equation:

$$\text{Normalized potency per 1 unit of exposure} = \sqrt{\frac{^2}{[\text{mass}_{\text{grams}}]^2}},$$

where $_2$ is the Potency coefficient from the MSW time-to-tumor modeling and mass is grams of asphalt fraction painted per application in the dermal bioassay.

The $_2$ term is designed to be multiplied by the square of the mass of the dose (in keeping with the quadratic shape of the dose–response curve, and to maintain parallelism with the quadratic dose–response for B[*a*]P). Figure 4 presents the relationship between the normalized potencies and the fluorescence measures. We adjusted this relationship by removing the lowest fractions with no carcinogenic response (Fractions D and E) and low EU/g measurements. This adjustment resulted in a linear relationship that describes the low-dose region of the existing data, where it will be important to apply this relationship to other asphalt mixtures. BURA-fume-fraction samples (Fractions D and E) were negative with regard to tumor induction and had fluorescence values less than 130. In addition, the paving asphalt examined by Clark et al. (2011) also exhibited no significant tumor induction (fluorescence value of 30 EU/g). Thus, existing data for asphalt samples below 130 EU/g do not show them to be related to an increase in skin tumors, and because the fractionation analysis (''Dose–response for asphalt mixtures'' section) indicated that the tumor response acts in an additive fashion, we would not expect a synergistic response in this concentration range. We therefore truncate our regression at 130 EU/g, to avoid over-predicting potency at very low concentrations.

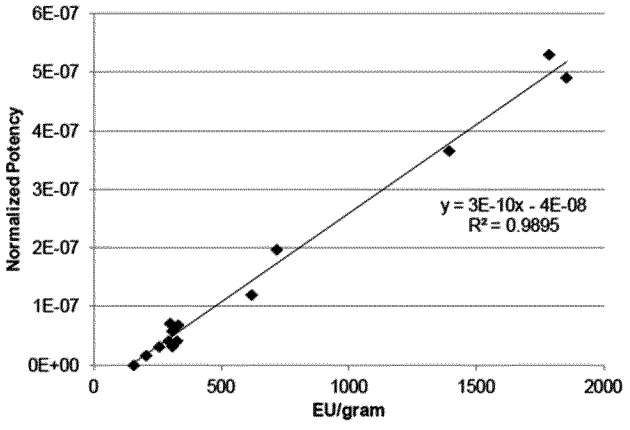

Figure 4. Correlation between fluorescence and normalized potency from MSW time-to-tumor modeling.

As shown in Table 15, predicted normalized potencies from the linear relationship are similar to the measured values. In addition, similar potencies are predicted for the whole asphalt fume (from Sivak et al. 1997), reconstructed fume (Fractions A + B+C + D+E), and the comparable laboratory-generated fume (LR-A) (from Clark et al. 2011).

The relationship between MSW-generated potencies and fluorescence could be used to predict the potential carcinogenic potency of a mixture based on its EU/g relative to the standard (laboratory-generated Type III whole asphalt fume). One would proceed by using the equation of the fitted line to calculate the normalized potency of the new mixture and adjust any carcinogenic potency by applying the ratio of this normalized potency to the normalized potency of the standard mixture. Thus, our linear regression (Figure 4) can be used to estimate the modeled carcinogenic potency based on fluorescence measures:

$$\text{Predicted normalized potency per 1 unit of exposure}(y) = 3 \quad 10^{10}(x) \quad 4 \quad 10^{8},$$

where "*x*" is the mixture's fluorescence (EU/g) and "*y*" is its predicted normalized potency. The standard of comparison – the laboratory-generated Type III whole asphalt fume – has an EU/g of 326, so its predicted normalized potency is $3 \quad 10^{10}$ (326) $4 \quad 10^{8} = 5.78 \quad 10^{8}$. An asphalt mixture with an EU/g of 1000 results in a predicted normalized potency of $3 \quad 10^{10}$ (1000) $4 \quad 10^{8} = 2.6 \quad 10^{7}$. This result can be related to the potency of the original whole asphalt laboratory-generated fume by a simple ratio: $2.6 \quad 10^{7} / 5.78 \quad 10^{8} = 4.5$ (i.e. the 1000 EU/g asphalt is 4.5 times as potent as the original laboratory fume). Thus, one could adjust the DSF of the whole fume ($1.46 \quad 10^{4}$ per μg/day in Table 10) by a factor of 4.5 to obtain a relative potency-adjusted DSF of $6.6 \quad 10^{4}$ per μg/day for the hypothetical asphalt mixture with a UV measure of 1000 EU/g.

However, this relationship has a limit below 134 EU/g, where potency cannot be quantified. Although we expected asphalts with fluorescence measures below this limit to have relatively low potency, it is likely greater than zero but not within the quantifiable range of this linear regression. Because we defined the MSW dose–response curves as parallel for the various tested mixtures, the relative potency calculated from the regression should apply to all parts of the dose–response relationship, for whatever dose level or exposure duration is calculated using the fitted MSW curves. As noted previously, to create unit risks for cancer risk calculations, we linearized the MSW's curvilinear dose–response below a defined POD, and the relative potencies calculated using the fluorescence adjustment apply to the risk calculations using unit risks as well. That is, when applied to a mixture with a given EU/g, the value of the unit risk is scaled up or down based on the ratio of predicted normalized potencies for the mixture of concern and for the standard mixture on which the unit risk derivation was based. This method can be used to qualitatively predict the potency of other similar asphalt mixtures, assuming that future tested mixtures have characteristics (e.g. PAH composition, distillation range and boiling points) similar to those of the original tested roofing and paving fumes. The relationship between fluorescence and carcinogenicity is based on the mouse skin painting studies; thus, extending the relationship to other exposures (e.g. inhalation) or endpoints (e.g. lung cancer) assumes that relative potency is conserved across exposure routes.

Inhalation carcinogenicity potency estimation

Because there is no animal cancer bioassay of BURA-fume inhalation, we examined inhalation cancer potency using a "sufficiently similar" mixture; i.e. we analyzed the results reported by Fuhst et al. (2007) for paving asphalt fumes. We also considered several relative potency approaches to leverage the skin-painting and paving fume results (Table 10).

Whole-mixture assessment

The paving asphalt bioassay performed by Fuhst et al. (2007) provides a mixture reasonably similar to BURA and is the key study examined in the whole-mixture inhalation assessment. The chemical composition of paving fumes was shown (in the "Asphalt chemistry and composition" section) to have a composition relatively similar to that of BURA fumes, save for lower concentrations of 4- to 6-ring PAHs (Table 4). Thus, we used the results of the paving fume study to give an indication of the potency of BURA fumes. As described previously, the animals examined in this bioassay did not experience a significant tumor response following exposure to asphalt fumes. We examined the dose–response for tumor incidence to estimate an upper-bound for estimate of cancer risks. Following EPA's (2012) recommended benchmark-dose–response procedures, we fit the inhalation data to the multi-stage cancer model and found a significant fit using the multistage-cancer model (3^{χ}). The estimated BMD and BMDL (at a benchmark response of 10%) were 300 and 195 mg/m^3 (total hydrocarbon content), respectively (Table 16). The BMDL is equivalent to 146.15 mg/m^3 adjusted from a

Table 16. Inhalation cancer potency estimates for alternative dose–response assessment methods.

Mixture	Critical study	Point of departure	Modeling approach	Potency estimate (IUR)	Relative potency ratio
Whole mixture assessment					
B[*a*]P	Thyssen et al. (1981, as reported by EPA, 2014a)	BMC_{10}=0.254 mg/m^3 $BMCL_{10}$=0.163 mg/m^3	MSW time-to-tumor model	6 10 ^{4} per μg/m^3	–
B[*a*]P	Heinrich et al. (1994)	$BMCL_{10}$=0.005 mg/m^3	MSW time-to-tumor model	2 10 ^{2} per μg/m^3	–
Paving asphalt	Fuhst et al. (2007)	BMC_{10}=300 mg/m^3 $BMCL_{10}$=195 mg/m^3 Total hydrocarbon content	Benchmark dose (Multistage-Cancer 3 [a])	6.84 10 ^{7} per μg/m^3 (adjusted to an 8-hour exposure)	(1.14 10 ^{3})[a] (3.42 10 ^{5})[a]
Relative potency assessment					
Total PAHs from lab-generated type III BURA fumes	Sivak et al. (1997)	IUR = 6 10 ^{4} per μg/m^3	EPA B[*a*]P IUR adjusted for relative potency of roofing asphalt to B[*a*]P	(5.2 10 ^{8} per μg/m^3)[b]	(8.6 10 ^{5})[c]
Total PAHs from field-generated BURA fumes (TR-A)	Clark et al. (2011)	IUR = 6 10 ^{4} per μg/m^3	EPA B[*a*]P IUR adjusted for relative potency of roofing asphalt to B[*a*]P	(1.1 10 ^{8} per μg/m^3)[b]	(1.8 10 ^{5})[c]
Total PAHs from lab-generated BURA fumes (sample LR-A)	Clark et al. (2011)	IUR = 6 10 ^{4} per μg/m^3	EPA B[*a*]P IUR adjusted for relative potency of roofing asphalt to B[*a*]P	(3.9 10 ^{8} per μg/m^3)[b]	(6.5 10 ^{5})[c]

B[*a*]P, benzo[*a*]pyrene; BURA, built-up roofing asphalt; EPA, United States Environmental Protection Agency; IUR, inhalation unit risk; LR, lab roofing; MSW, multistage-Weibull; PAH, polycyclic aromatic hydrocarbons; TR, tank roofing.

[a]Relative potency estimated as the ratio of the IUR for asphalt mixtures to the IURs from either dose–response assessment for B[*a*]P.

[b]The IURs were estimated by multiplying EPA's draft IUR (6 10 ^{4} per μg/m^3) by the concurrent RPF.

[c]Relative potency estimates based upon the results from the skin painting carcinogenicity bioassays (Table 10).

6-h exposure to an 8-h work day exposure) as the POD for estimating relative cancer risks based on paving asphalt fume. The resulting inhalation unit risk (IUR) is estimated (0.1/146.25 mg/m^3) at 6.84 10 ^{4} per mg/m^3.

By comparison, the potency of the asphalt fume is 3–5 orders of magnitude less than that of B[*a*]P (Table 16). We consider this IUR for paving asphalt to be a mixture sufficiently similar to BURA asphalt fumes and have used it in the "Cancer risk estimates for roofing workers" section to estimate cancer risks for roofing workers exposed to BURA fumes.

Relative potency assessment

The potency of BURA fumes relative to B[*a*]P can be assessed using two methods, depending on the source of information about potency of fumes relative to that of B[*a*]P. In the first method, we assumed that BURA asphalt fumes are generally similar in composition to asphalt paving fumes and then determined a relative inhalation potency for asphalt fumes compared to B[*a*]P, using the observed potency of asphalt fumes for lung cancer (as determined in the "Whole-mixture assessment" section using the Fuhst et al. 2007 study) and an inhalation potency for B[*a*]P estimated from a parallel study by Heinrich et al. (1994). The Fuhst et al. (2007) study of paving fume inhalation did not include a concurrently tested B[*a*]P positive control, but the authors of this study cited the earlier Heinrich et al. study as an effective positive control, stating that their study was intended to replicate the methods of Heinrich et al. (1994).

In the Heinrich et al. (1994) study, female Wistar rats inhaled a cooking gas that contained 20 or 46 μg/m^3 B[*a*]P over a 5 day per week, 17 h per day period of 10 and 20 months, followed by exposure to clean air for 10 or 20 months. Lung tumors (primarily squamous cell carcinomas) were reported at rates of 38.9% and 97.2% for the 20 or 46 μg/m^3 B[*a*]P groups, with no comparable lesions in the control group. The IUR developed by Heinrich et al. using the multistage model is 2 10 ^{2} per μg/m^3. The quantitative potency estimates for B[*a*]P are summarized in Table 16. The relative inhalation potency of paving fumes (as determined from our analysis of Fuhst et al. 2007) to that of B[*a*]P (as determined from our analysis of Heinrich et al. 1994) is 3.42 10 ^{5}. This number can be used to scale amounts of BURA fume inhalation into inhaled B[*a*]P equivalents, and then human risks can be projected using an externally supplied human inhalation potency for B[*a*]P, such as the EPA's recent draft value (EPA 2014a).

This EPA draft value is based on carcinogenicity observed (for tumors in the larynx, pharynx, trachea, esophagus and forestomach) in male Syrian golden

hamsters exposed to B[*a*]P condensed onto sodium chloride (NaCl) particles following chronic inhalation exposure (Thyssen et al. 1981, as cited in EPA 2014a). The draft IUR of 6 10 ^{4} per μg/m^{3} was calculated using the MSW time-to-tumor model and linear extrapolation from a $BMCL_{10}$ of 0.163 mg/m^{3} for the occurrence of upper respiratory tract and pharynx tumors (all treated as incidental death). EPA (2014a) recognized that the results from Thyssen et al. (1981) found cancers of the upper respiratory tract but not lung tumors and, so, may not provide an ideal study for extrapolating to humans. Nonetheless, EPA provides the B[*a*]P IUR as a basis for estimating general cancer risks from inhaled fumes, and this IUR is expected to serve as the basis for estimating inhalation cancer risks for other PAHs using their potency relative to B[*a*]P. For the purposes of this assessment, we are using EPA's potency estimate for B[*a*]P as an indicator of potential risks of respiratory cancers, recognizing that species extrapolation from these data involves potential uncertainties.

The foregoing approach to determining the relative inhalation potency of asphalt fumes and B[*a*]P has some notable but unavoidable shortcomings: as noted, the Fuhst et al. (2007) study is of paving asphalt fumes, not BURA fumes; it also showed no significant response, and the fume potency is determined as an upper bound; the Heinrich et al. (1994) study is not a true concurrent positive control, and moreover the exposure was not to pure B[*a*]P but to an aerosol produced by heating coal tar, with the effects attributed by the study's authors to its B[*a*]P content.

Accordingly, we considered an alternative approach for defining relative inhalation potency: using the relative potency of BURA and B[*a*]P that we determined from the skin-painting studies of Sivak et al. (1997), making the assumption that relative potency is conserved across exposure routes. For example, we adjusted the draft EPA B[*a*]P IUR of 6 10 ^{4} per μg/m^{3} by a potency ratio of 8.6 10 ^{5} (based on the laboratory-generated asphalt fume in the skin-painting assay conducted by Sivak et al. compared to its B[*a*]P concurrent positive control), resulting in an adjusted IUR of 5.2 10 ^{8} per μg/m^{3} (as shown in Table 16). There are reasons to question whether the relative potency is indeed independent of route of exposure and target tissue, because absorption and local metabolic activation may differ, as may inherent vulnerabilities of tissues. Nonetheless, this approach provides a window into possible fume potency that results from observations of tumorigenicity of BURA fumes, rather than paving asphalt fumes.

Both of these approaches represent less-than-ideal comparisons, necessitated by the nature of the studies available. But in fact, the two approaches do not disagree markedly with one another (i.e. the IURs from the whole paving fume is near the range of IURs estimated by using the RPFs from the skin-painting studies; see Table 16). We further used the IURs generated by adjusting B[*a*]P B[*a*]P to laboratory- or field-generated relative potency ratios to estimate risks for roofing workers (see the "Cancer risk estimates for roofing workers" section).

Cancer risk estimates for roofing workers

We considered a number of alternative methods for characterizing the carcinogenic potency of BURA via inhalation or dermal exposure in the dose–response analysis. Through extensive modeling and analysis, we developed several quantitative estimates of risk criteria (e.g. lope factors or RPFs) and combined them with exposure information to estimate risks to roofing workers. Each of our analysis methods (i.e. whole-mixture, mixture-components, and relative potency) had unique challenges owing to the differing test methods, exposure conditions and data types. Further, each of these methods carries a number of assumptions that may weaken the reliability of any one method. Nevertheless, the disparate data sets enabled alternative quantitative comparisons, allowing several lines of evidence to be combined to develop broad conclusions regarding the risks from exposure to BURA. Therefore, we explored a range of risk characterization methods based on the quantitative risk criteria developed in the previous section.

Inhalation risk assessment

Lung cancer risk is the endpoint most suggestive of association with roofer exposure in epidemiological studies. The available animal bioassay data on inhalation do not show tumor elevation, but lung cancer is the most pronounced effect, and it is the basis of our upper-bound estimate of potential cancer risks from fume inhalation, as described above. Accordingly, we used projected potential lung cancer risk as a measure of total cancer risk to roofers from inhalation. We combined estimates of inhalation exposures among roofers with IURs as developed above from the rat bioassay data. As we noted during the derivation of this IUR, the available basis for estimating it is meager, and some strong assumptions about the validity of approximations are necessary. This approach, despite its shortcomings, nonetheless provides the best available means to project a possible inhalation cancer risk for roofing workers,

though the limitations in estimating these risks should be borne in mind.

We developed a generalized occupational exposure estimate for roofing workers to evaluate lung cancer risks from exposure to asphalt fumes. Occupational risk assessments generally assume long-term exposure of adults for a typical 40-h workweek, 250 days/year, for a duration of 45 years. We note that EPA (1989, 2004) typically uses a 30-year duration for exposure assessments; thus, the aforementioned exposure model reflects a greater exposure level (i.e. more conservative) than would be estimated following EPA risk assessment methods. Asphalt roofing work is seasonal in some portions of the USA, and the working tenure is typically shorter than 30 years (see the "Occupational exposure" section). Nevertheless, this general exposure model provides a consistent and conservative estimate for characterizing potential risks. Inhalation exposure is quantified by the following general equation:

$$\mathrm{ADD_{air}} = \frac{(C_{\mathrm{air}} \quad \mathrm{ET} \quad \mathrm{EF} \quad \mathrm{ED})}{\mathrm{AT_c}},$$

where $\mathrm{ADD_{air}}$ is the occupational average daily dose (ADD, μg/m^3); C_{air} is the concentration of BURA in air (μg/m^3); ET is exposure time (8 working hours/day); EF is the exposure frequency (250 working days/year); ED is the exposure duration (45 year working duration) and $\mathrm{AT_c}$ is the averaging time over a lifetime (365 days/year 70 years 24 h/day).

We reviewed typical air concentrations for roofing workers in the "Occupational exposure" section. We conducted risk calculations using GM-measured concentrations from asphalt samples and assumed that asphalt emissions were kept at or below the current TLV of 0.5 mg BSF/m^3. We assumed that PAHs contribute most of the biologically active compounds within asphalt emissions. Kriech et al. (2007) provided PAH estimates of field exposures during roofing operations (Table 5). Total PAH concentrations and total B[*a*]P-equivalent concentrations are presented (normalized to the TLV limit) in Table 17.

Risk estimates based on the inhalation ADDs and varying IURs and RPFs are discussed in the following sections. Given that the most widely accepted application of an RPF-based approach involves evaluating mixtures of PAHs, based on the potency of the PAH B[*a*]P, we have applied (for comparison purposes) the EPA (1993) PAH RPF approach to the PAH content of asphalt fumes. We explored other sets of RPFs (results not presented) and the risk estimates were generally consistent within an order of magnitude depending on the set of RPFs evaluated (e.g. EPA 2010 or WHO).

We calculated inhalation cancer risk estimates using either total PAH or B[*a*]P-equivalent exposure concentrations and compared them to various IURs (Table 17). All combinations yielded relatively low cancer risk estimates (i.e. within typically accepted risk levels of 10^{-4} and 10^{-6}). Estimates based on the B[*a*]P-equivalent approaches resulted in higher cancer risk estimates. This finding may indicate that the RPF approach for inhalation exposures over-estimates potential risks. Risk estimates based on paving fumes or relative potency to B[*a*]P yielded low risk estimates within a factor of 10, and laboratory-generated fumes resulted in slightly higher risk estimates. Our occupational risk assessment, using a range of alternative exposure and toxicity estimation approaches (each with their own advantages and disadvantages), resulted in a consistent and low pattern of estimated risks. This quantitative analysis is also consistent with the epidemiology data for lung cancer (and possibly upper respiratory tract cancers), which show a weak and uncertain relationship for roofing workers. Based on these results, the current TLV for asphalt emissions of 0.5 BSF/m^3 is adequately protective for inhalation exposures experienced by roofing workers.

Dermal risk assessment

Epidemiological studies show little evidence of elevated skin cancer risks among roofers. Nonetheless, we combined estimates of dermal exposure in roofers with the skin-painting animal-bioassay analyses above, for which the observed effect was skin cancer, to determine a human-applicable projection of potential cancer risks for roofers from dermal exposure. Skin-painting bioassays did not observe elevation of internal tumors, but our analysis cannot directly address the possibility that systemically absorbed material from dermal exposure could affect risks of internal cancers to some degree.

We calculated dermal exposure and potential risks to asphalt fume condensates using a generalized occupational exposure estimate. We used a typical worker exposure model to estimate the ADD for dermal exposure, and it is quantified by the following general equation:

$$\mathrm{ADD_{skin}} = \frac{(C_{\mathrm{skin}} \quad \mathrm{FA_{skin}} \quad \mathrm{SA} \quad \mathrm{EF} \quad \mathrm{ED})}{\mathrm{AT_c}},$$

where $\mathrm{ADD_{skin}}$ is the occupational average daily dose (μg/day); C_{skin} is the concentration of BURA on skin (μg/cm^2 day); $\mathrm{FA_{skin}}$ is the fraction of asphalt constituents absorbed by skin (%); SA is the surface area of exposed skin (3300 cm^2 of exposed skin); EF is the exposure frequency (250 working days/year); ED is the exposure

Table 17. Occupational inhalation exposure and risk estimates for BURA fumes.

Exposure metric	Asphalt fume concentration[a] (mg/kg TOM)	Exposure estimate ($\mu g/m^3$)[b,c]	ADD ($\mu g/m^3$)[d]	IUR ($\mu g/m^3$)$^{-1}$	IUR source[e]	Inhalation cancer risk estimate
Field-generated fumes						
Total PAHs (mean)	603	1.34	0.20	6.84E 07	Paving Asphalt	1E 07
Total PAHs (maximum)	603	6.03	0.88	6.84E 07	Paving Asphalt	6E 07
Total PAHs (mean)	603	1.34	0.20	1.10E 08	EPA IUR Adjusted by RPF (TR-A)	2E 09
Total PAHs (maximum)	603	6.03	0.88	1.10E 08	EPA IUR Adjusted by RPF (TR-A)	1E 08
B[*a*]P_{eq} (mean)[f]	5.38	0.01	0.002	6.00E 04	EPA Draft IUR	1E 06
B[*a*]P B[*a*]P_{eq} (maximum)	5.38	0.05	0.01	6.00E 04	EPA Draft IUR	5E 06
Lab-generated fumes						
Total PAHs (mean)	672	1.49	0.22	7.73E 07	Paving Asphalt	2E 07
Total PAHs (maximum)	672	6.72	0.99	7.73E 07	Paving Asphalt	8E 07
Total PAHs (mean)	672	1.49	0.22	5.20E 08	EPA IUR Adjusted by RPF (Type III)	1E 08
Total PAHs (maximum)	672	6.72	0.99	5.20E 08	EPA IUR Adjusted by RPF (Type III)	5E 08
Total PAHs (mean)	672	1.49	0.22	3.90E 08	EPA IUR Adjusted by RPF (LR-A)	9E 09
Total PAHs (maximum)	672	6.72	0.99	3.90E 08	EPA IUR Adjusted by RPF (LR-A)	4E 08
B[*a*]P_{eq} (mean)	8.52	0.02	0.003	6.00E 04	EPA Draft IUR	2E 06
B[*a*]P_{eq} (maximum)	8.52	0.09	0.01	6.00E 04	EPA Draft IUR	8E 06

ADD, average daily dose; B[*a*]P_{eq}, benzo[*a*]pyrene equivalent concentration; BSF, benzene-soluble particulate fraction; BURA, built-up roofing asphalt; EPA, United States Environmental Protection Agency; IUR, inhalation unit risk; LR, lab roofing; PAH, polycyclic aromatic hydrocarbons; TOM, total organic matter; TR, tank roofing.

[a]Arithmetic mean (AM) concentrations in mg/kg TOM calculated for roofing field samples (TR-A, TR-B, TR-C, TR-D) or roofing lab sample (LR-A) (Table 5) as reported by Kriech et al. (2007).

[b]Normalized air concentration calculated as follows: conc. (mg/kg TOM) ratio of TOM/BSF 0.5 mg BSF/m^3 1000 μg/mg 10 ^{6} kg/mg.

[c]The mean and upper bound TOM/BSF ratios are 4.44 and 20 developed from paired samples collected by Kriech et al. (2004a).

[d]ADD calculated as described in the "Inhalation risk assessment" section.

[e]IURs defined in Table 14.

[f]All B[*a*]P equivalent concentrations estimated using EPA's (1993) RPFs.

duration (45 year working duration) and AT_c is the averaging time over a lifetime (365 days/year 70 years).

The most robust assessment of dermal exposure was reported by McClean et al. (2007), who measured PAH exposures during roof tear-off, put-down and kettle operations under typical roofing scenarios (see discussion in the "Occupational exposure" section). McClean et al. (2007) evaluated exposure during both roof removal and roof laying, while controlling for exposure to coal tar pitch. Based on an evaluation of task-based exposures on non-pitch days, the best estimates of dermal exposures related to roofing asphalt were represented by the adjusted mean PAH concentrations of 340 ng/cm^2 for tear-off, 161 ng/cm^2 for put-down and 77 ng/cm^2 for kettlemen (McClean et al. 2007). Total PAH concentrations are assumed to represent the biologically active components of BURA asphalt vapors and aerosols.

Dermal *in vitro* studies using human skin exposed to asphalt condensates indicate a slower absorption rate (dermal flux) than neat B[*a*]P solutions (such as the dosing regimen used to evaluate dose–response of skin tumors in mice by Sivak et al. 1997) (see the "Occupational exposure" section). In addition, dermal penetration and metabolism of PAHs within the epidermal layer is similarly low in human skin compared to mice (see the "Occupational exposure" section). Based on this information, we examined a range of absorption fractions of the biologically active PAHs in the whole BURA fume condensate that would pass through the stratum corneum to the epidermal layer and be available to react with the target skin cells (i.e. keratinocytes). The dermal risk assessment results are presented in Tables 18 and 19, assuming a 1% absorption fraction for PAHs in BURA fume condensates. This semi-quantitative absorption fraction is intended to reflect differences in species toxicokinetics (not addressed by the body weight scaling approach performed in the "Dose–response assessment methods" section), differences in species and mixture specific dermal absorption rates (discussed in the "Occupational dermal exposure" section), and differences in occupational exposures as compared to the laboratory (e.g. routine bathing between shifts versus 24-h constant dosing in the laboratory mice). Risk estimates change proportionally with assumed absorption fraction (i.e. lower absorption reduces risk estimates and *vice versa*). We note that dermal absorption is a key uncertainty in the risk assessment, as there is no established method for relating the absorption and metabolism of PAHs from mouse skin to human skin and to account for differing exposure conditions from the

Table 18. Occupational dermal exposure and risk estimates for BURA fume condensates (without coal tar present).

Worker type	Dermal exposure estimate (μg PAHs/cm^2-day)[a]	ADD (μg/day)[b]	DSF (μg/day) $^{-1}$	DSF source[c]	Skin cancer risk estimate
Roofing worker[d]	0.501	7.28	3.47 10 ^{5}	Field Fume (TR-A)	3E 04
Tear-off	0.340	4.94	3.47 10 ^{5}	Field Fume (TR-A)	2E 04
Put-down	0.161	2.34	3.47 10 ^{5}	Field Fume (TR-A)	8E 05
Kettleman	0.077	1.12	3.47 10 ^{5}	Field Fume (TR-A)	4E 05
Roofing worker[d]	0.501	7.28	1.46 10 ^{4}	Lab Fume (Type III)	1E 03
Tear-off	0.340	4.94	1.46 10 ^{4}	Lab Fume (Type III)	7E 04
Put-down	0.161	2.34	1.46 10 ^{4}	Lab Fume (Type III)	3E 04
Kettleman	0.077	1.12	1.46 10 ^{4}	Lab Fume (Type III)	2E 04
Roofing worker[d]	0.501	7.28	1.06 10 ^{4}	Lab Fume (LR-A)	8E 04
Tear-off	0.340	4.94	1.06 10 ^{4}	Lab Fume (LR-A)	5E 04
Put-down	0.161	2.34	1.06 10 ^{4}	Lab Fume (LR-A)	2E 04
Kettleman	0.077	1.12	1.06 10 ^{4}	Lab Fume (LR-A)	1E 04

ADD, average daily dose; BURA, built-up roofing asphalt; DSF, dermal slope factor; LR, lab roofing; PAH, polycyclic aromatic hydrocarbon; TR, tank roofing.
[a]Dermal exposure estimates based on total PAHs (with and without coal tar exposure) from McClean et al. (2007).
[b]ADD calculated as described in the "Dermal risk assessment" section, assuming a 1% absorption of PAHs in BURA fume condensates.
[c]DSFs defined in Table 10.
[d]Note that the roof worker data reflects a full shift exposure (tear-off plus put-down), while the tear-off and put-down data represent half-shift exposures.

Table 19. Occupational dermal exposure and risk estimates for asphalt fume condensates with coal tar present.

Worker type	Dermal exposure estimate (μg PAHs/cm^2 day)[a]	ADD (μg/day)[b]	DSF (μg/day) $^{-1}$	DSF source[c]	Skin cancer risk estimate
Roofing worker[d]	0.898	13.05	3.47 10 ^{5}	Field Fume (TR-A)	5E 04
Tear-off	0.886	12.87	3.47 10 ^{5}	Field Fume (TR-A)	4E 04
Put-down	0.344	5.00	3.47 10 ^{5}	Field Fume (TR-A)	2E 04
Kettleman	0.299	4.34	3.47 10 ^{5}	Field Fume (TR-A)	2E 04
Roofing worker[d]	0.898	13.05	1.46 10 ^{4}	Lab Fume (Type III)	2E 03
Tear-off	0.886	12.87	1.46 10 ^{4}	Lab Fume (Type III)	2E 03
Put-down	0.344	5.00	1.46 10 ^{4}	Lab Fume (Type III)	7E 04
Kettleman	0.299	4.34	1.46 10 ^{4}	Lab Fume (Type III)	6E 04
Roofing worker[d]	0.898	13.05	1.06 10 ^{4}	Lab Fume (LR-A)	1E 03
Tear-off	0.886	12.87	1.06 10 ^{4}	Lab Fume (LR-A)	1E 03
Put-down	0.344	5.00	1.06 10 ^{4}	Lab Fume (LR-A)	5E 04
Kettleman	0.299	4.34	1.06 10 ^{4}	Lab Fume (LR-A)	5E 04

ADD, average daily dose; BURA, built-up roofing asphalt; DSF, dermal slope factor; LR, lab roofing; PAH, polycyclic aromatic hydrocarbon; TR, tank roofing.
[a]Dermal exposure estimates based on total PAHs (with and without coal tar exposure) from McClean et al. (2007).
[b]ADD calculated as described in the "Dermal risk assessment" section, assuming a 1% absorption of PAHs in asphalt fume condensates.
[c]DSFs defined in Table 10.
[d]Note that the roof worker data reflects a full shift exposure (tear-off plus put-down), while the tear-off and put-down data represent half-shift exposures.

mouse bioassay to the field situation. Further, we have no quantitative basis to estimate systemic effects from dermal exposure. Thus, the dermal assessment assumes that the risk is principally dermal cancer and does not address the possibility of further risks from PAHs that are absorbed systemically via dermal exposure. Therefore, the results of the dermal risk assessment should be considered rough estimates.

We assumed that clothing and personal protective equipment would cover most of roofing workers' skin, while portions of the head, hands and forearms may be exposed during BURA applications. Further, current personal hygiene practices would limit prolonged dermal exposure for workers, whereas the mice in the skin-painting assays were exposed to increasing mixture applications without washing between doses. Per EPA (2004) recommendations, we assumed a median skin surface area for males and females for the head, forearms and hands that was available for exposure to BURA vapors and aerosols during a typical work day (approximately 3300 cm^2).

Based on the generalized exposure model, we considered several alternative exposure metrics to assess the potential skin cancer risk from exposure to BURA fume condensates (Table 18). The field-based (TR-A BURA fume) cancer risks estimates (at 1% absorption) were within the range of 4 10 ^{5} to 3 10 ^{4} when coal tar exposures were excluded (Table 18). Risk estimates using potency estimates from laboratory-generated fume studies resulted in somewhat higher risks (Table 18). Workers conducting roofing tear-off and/or put-down may also be exposed to coal tar (owing to historical use of coal tar in roofing materials and, hence, higher exposure to carcinogenic PAHs than would occur during modern use of BURA), and this would yield higher estimates of cancer risk depending on the slope factor applied (as shown in Table 19). Although available data suggest that dermal absorption is limited (see the

"Occupational exposure" section), if a higher rate of absorption (e.g. 10%) is considered, then estimated risks increase proportionately. Thus, dermal absorption is a key variable when considering skin cancer risks. We have assumed constant dermal exposure throughout the duration of a work-life (45 years), which is likely an overestimate based on current worker demographics (see the "Occupational exposure" section). This suggests that our quantitative approach yields conservative (and perhaps overestimates) dermal cancer risks as these findings are consistent with results from epidemiology studies, indicating no significant association between roofing workers' exposures and the development of skin cancer (see the "Epidemiology data" section).

Uncertainty analysis

We conducted an uncertainty analysis to determine how the various assumptions and data gaps identified for the various studies and modeling approaches affected results (EPA 2000). For each of the preceding sections, data limitations or uncertainties are discussed to document any biases in the overall risk assessment. In some cases (e.g. MSW time-to-tumor modeling), a quantitative assessment of uncertainty in some of the input parameters is accounted for in the modeling process in the generated confidence levels. The following sections summarize other sources of uncertainty that should be considered when evaluating the risk assessment results.

BURA chemical profile

A key component of risk assessment is the characterization of the mixture under study, or sufficiently similar mixtures, to accurately relate exposure and effect (EPA 2000). As described in the "Asphalt chemistry and composition" section, a number of studies have provided detailed characterizations of BURA and other asphalts under varying environmental conditions. The variability of composition of components appears to be limited based on current information, suggesting that one can extrapolate results from tested BURA samples to non-tested BURA or similar asphalts. Further, the data are robust for characterizing carcinogenic PAHs within BURA and the changes in their composition and amounts during environmental transformations (e.g. high-temperature heating). The primary uncertainty associated with BURA chemical profiles lies with uncharacterized PAHs (e.g. heterocyclic PAHs). Most studies have focused on carcinogenic PAHs, but a number of different classes of PAHs are found in BURA and other asphalts. The extent to which these uncharacterized compounds contribute to the overall exposure and risks is unknown. However, our analysis for whole asphalt fumes (whole-mixture) using various methods (e.g. MSW modeling or fluorescence measures) incorporates all PAH classes and therefore, any additive, competitive or synergistic actions were captured in these analyses. Thus, this uncertainty is tempered, to a degree, by considering the whole asphalt mixture.

Exposure estimation

The existing exposure information is suitable for characterizing general exposure levels during various roofing activities (see discussion in the "Occupational exposure" section). However, interpretation of these results is hampered by varying analytical methodologies, confounding exposures to other contaminant sources (e.g. coal tar), and/or limited chemical characterizations (i.e. mall number of PAH concentrations reported). In some cases, environmental concentrations of PAHs may be biased high (i.e. if coal tar is present) or biased low (if a full characterization of the chemical profile is not provided). Several investigators have attempted to address these issues by incorporating alternative measurement techniques or investigation methods (e.g. Kriech et al. 2004a; McClean et al. 2007). This approach is particularly relevant to the use of RPFs for PAHs, since in some cases, existing exposure data do not provide full concentration profiles for the PAHs of concern. Thus, applying an RPF approach is complicated to a degree by limitations in existing data.

Epidemiology studies are insufficient to characterize the relationship between BURA exposure and skin or lung cancer incidence. As described in the "Epidemiology data" section, numerous limitations in the current literature are recognized, including a lack of exposure data, exposure misclassifications, and confounding contaminant exposures (e.g. coal tar, asbestos, tobacco smoke). Therefore, the epidemiology information provides an unreliable source for exposure information, and it has limited utility for quantitatively characterizing risks (as discussed in the "Epidemiology data" section).

Finally, as discussed in the "Occupational exposure" section, application of roofing asphalt is typically a small percentage of total roofing hours for roofing workers (ARMA et al. 2011). Further, specialization of workers (by job type) has decreased, and thus a typical worker is not expected to spend a full work shift installing BURA throughout their work tenure. However, in our risk assessment ("Cancer risk estimates for roofing workers" section), we have conservatively assumed that a worker is exposed to BURA fumes over the working lifetime.

This represents an upper bound assumption and likely overestimates exposure and risk.

Dermal absorption

A key consideration for evaluating skin exposure and skin tumor development from BURA includes an assessment of dermal absorption. Information on the dermal absorption of individual PAHs indicates that PAHs permeate through animal or human skin with differences occurring at differing anatomical sites or from differing exposure matrices (e.g. solvents, soils) (ATSDR 1995; EPA 2014b). Dermal absorption of asphalt fume condensates (*in vitro* human skin studies) is described in the "Occupational exposure" section and indicates that a small portion of the skin-applied dose crosses the stratum corneum (or the rate of absorption is relatively slow as compared to pure B[*a*]P) and is available to systemic circulation. However, whether dermal absorption of BURA fume condensates is similar for both human skin and mouse skin (as examined in the skin-painting bioassays) is uncertain. Absorption studies for B[*a*]P demonstrate that mouse skin may absorb this compound to a higher degree than human skin does (EPA 2014a; Kao et al. 1985; Storm et al. 1990). For example, Storm et al. (1990) compared B[*a*]P absorption across species and found that human skin absorbed 31% of the applied dose, whereas mouse and guinea pig skin absorbed 55–60% of the applied dose. Further, dose levels in mouse skin-painting studies were magnitudes higher than those typically encountered in the work environment. Therefore, it is unknown to what degree much higher doses lead to higher rates of absorption and thus increased tumor incidence, and whether this skin tumor incidence is directly related to roofing workers. Assuming that either skin absorption of BURA fume condensates in mice is equivalent to absorption rates in humans or that BURA fume condensates are absorbed at the same rate as an indicator substance (such as B[*a*]P) may bias risk estimates high if true absorption rates are limited (as demonstrated in the *in vitro* studies).

DSF endpoint

The DSFs estimated in the "Quantitative dose–response assessment" section are derived to specifically address the probability of skin cancers. Although a small fraction of PAHs from asphalts may be absorbed systemically (see the "Occupational dermal exposure" section), most of the material is either not absorbed or is retained within the viable layers of the epidermis. At present, we have no data to quantify the probability of tumor incidence at distal sites following dermal exposure. EPA (2014a) also recognized this uncertainty in their draft DSF analysis and suggested that existing data support the conclusion that the risk of skin cancer likely outweighs cancer risks at other sites. Therefore, this analysis is unable to estimate cancer risks other than those for skin or lung tumors. Because the available data are on skin cancers, our focus on them is consistent with our aim to evaluate the possible quantitative risk implications for roofing workers of those studies that form the basis for identifying asphalt fume exposures as potentially carcinogenic.

Dermal cross-species dose scaling procedures

We determined human equivalent DSFs from the PODs generated from the MSW dose–response modeling. We calculated the POD HEDs in a manner consistent with EPA's (2005) Cancer Guidelines, using an adjustment for cross-species scaling to address toxicological equivalence across species. The POD was converted to an HED on the basis of body weight (EPA 1992). This conversion was accomplished by multiplying the POD by the $(BW_{human}/BW_{Mouse})^{3/4}$ (EPA 1992), where the mouse and human body weights were EPA default reference body weights $(70\,kg/0.035\,kg)^{3/4}$. Different methodologies have been established for interspecies scaling of PODs used to derive oral slope factors. Cross-species adjustment of oral doses is based on allometric scaling using the ¾ power of body weight. This adjustment accounts for the relatively rapid distribution, metabolism, and clearance in small rodents as compared to larger humans (EPA 2005). No established method exists to adjust for interspecies differences in dermal toxicity at the point of contact; however, allometric scaling using body weight to the ¾ power has been recently evaluated by EPA (2014a, b) based on known species differences in dermal metabolism and penetration of B[*a*]P. We have reviewed these methods and agree that the approach selected by EPA is reasonable for addressing the extrapolation across species. However, we recognize that this adjustment may not account for differences with extrapolating mouse skin-painting studies to real world human exposures, such as:

> *Dosimetry of dermal exposure*: Skin-painting studies incorporate an exposure paradigm that continually adds the chemical without washing or removing unabsorbed fractions. Therefore, the mouse skin may accumulate a depot of chemical (i.e. applied PAHs are retained within the skin layers and not fully absorbed) that results in exposures greater than an average administered daily dose. Several authors

have reported this phenomenon (i.e. Kao et al. 1988; Moody et al. 2007). Therefore, the dosimetry of skin exposure is uncertain with regard to transforming skin-painting exposure to worker exposures. In our time-to-tumor modeling, any accumulative effect of repeated dosing on skin concentrations is accounted for in the model's estimate of the time dependence of tumor induction. In applying this model to human risk estimation, we are assuming that the time dependence under ongoing exposure applies as well, albeit scaled to a human timescale (as embodied in the cross-species dose scaling assumptions being applied). However, we presume that post-shift dermal deposition is washed off by roofing workers, limiting the dermal uptake in a way that does not apply to the mice.

Thickness of epidermal layer: The thickness of epidermal layers is known to be greater in human skin than mouse skin (e.g. 0.0104 mm in mouse skin vs. 0.052 mm in human skin) (see review by Knafla et al. 2011). This discrepancy may account for some of the differences observed in mice bioassays, while human epidemiological studies are largely negative.

Skin cells at risk: The default cross-species scaling method (discussed above) makes the implicit assumption that the number of cells at risk is proportional to body mass. In the case of skin, humans have many more cells at risk than do mice, but the proportionality is less than the difference in body mass would indicate. Brown et al. (1997) cite data indicating that human skin constitutes about 3.71% of body mass, while mouse skin constitutes an estimated 16.54%. The difference reflects the lower surface:volume ratio of humans compared to mice, partly tempered by the fact that human skin is thicker. An implicit organ-to-organ scaling based on mass would overestimate human risk by a factor of about 4.5 based on the difference between a body mass ratio of 2000 for a 70 kg human and a 35 g mouse, and a skin mass ratio of 449 (i.e. 2000/449 = 4.5). We did not adjust risk estimates for this difference; therefore, our dermal risk estimates may be overestimated.

Dermal metabolism: Species differences are apparent for skin PAH metabolism (e.g. Bronaugh et al. 1994; Knafla et al. 2011; Storm et al. 1990). For example, Storm et al. examined absorption and metabolic activity with B[*a*]P and found that activity of AHH and ethoxycoumarin deethylase was significantly lower in human skin than in mouse skin. B[*a*]P was also less well absorbed by human skin in this study (55–60% in mouse and guinea pig skin vs. 31% in human skin). Therefore, absorption rate and dermal metabolism may also be limiting factors for cross-species extrapolation. Our dermal risks likely are overestimated, because none of the aforementioned factors were incorporated into the risk assessment (as no standard method exists). Nevertheless, the direct potency estimates examined in the "Quantitative dose–response assessment" section among B[*a*]P and BURA in mice demonstrate that BURA is significantly less potent than B[*a*]P, and the potencies may be further reduced if cross-species extrapolation factors are considered.

Inherent sensitivity to carcinogenesis of mouse versus human skin: Not only do the epidemiological studies of roofers (reviewed above) show no skin cancer elevation, further studies of the therapeutic use of dermal application of coal tar to treat eczema fail to show skin tumor risks from even decades of dermal exposure to a PAH mixture that ought to be supposed to be more potent than asphalt fumes (Pittelkow et al. 1981). Human xenografts onto rodent skin subjected to skin painting studies with PAHs that are tumorigenic in mouse skin do not develop tumors, while the adjacent rodent skin does (Atillasoy et al. 1997; Soballe et al. 1996; Urano et al. 1995). PAH-induced mouse skin tumors have a genetic signature that differs from the genetic signature of human skin cancers (Balmain & Harris 2000). Our assumption that (after appropriate exposure scaling as discussed above) human skin is equally sensitive to tumor induction as mouse skin may be in error. There is no evident way to measure and account for a quantitative difference in inherent sensitivity, but to the extent one exists, our risk estimates for roofing workers' skin cancer risk from dermal exposure may be too high.

RPF approach: The EPA (1993) RPF approach for PAH-containing mixtures does not appear to be viable for BURA fumes (e.g. from the Clark et al. 2011; Fuhst et al. 2007; Sivak et al. 1997 studies). The observed potency of BURA fumes as a tested mixture is substantially greater than would be projected by considering the concentrations and RPFs of individual constituent PAHs that have RPF values available. This observation is consistent with previous findings indicating that PAH content is not an effective predictor of the carcinogenicity of BURA fumes (Sivak et al. 1997), and that other non-quantified PAHs may be contributing to the tumor response (Clark et al. 2011). One possible explanation is that that the more potent laboratory-generated BURA produced at high temperatures (e.g. 316 °C) does not appear to have an appreciably higher level of PAH content. For example, as the

temperature increases from 230–316 °C, the amount (mass) of fume emitted increases by a factor of 50–80, but the concentration of the PAHs in the emitted BURA fume does not scale similarly and increases only by two- to threefold. Another potential explanation is that the carcinogenic PAHs are present at very low levels in BURA. For example, the 4- to 6-ring PAHs, which are generally considered to be the most toxicologically relevant (EPA 2010), were either not detectable or present at concentrations below a biological effect level (Sivak et al. 1997). This profile is typical of all petroleum distillate-based chemical mixtures such as BURA, which typically contain more highly alkylated and heterocyclic congeners that contain sulfur, nitrogen and oxygen (Feuston et al. 1994; Roy et al. 1998; Sivak et al. 1997). Finally, the discordance between the expected and observed results (Table 14) may indicate that in this case the EPA RPF method (relying on a subset of PAHs) is not robust enough to accurately quantify the carcinogenic potential of BURA.

Route-to-route extrapolation: Because no inhalation bioassay directly measures BURA fumes, we evaluated the relative potency of BURA from skin-painting assays to adjust inhalation risks relative to B[*a*]P. This approach assumes that chemical exposure, toxicokinetics, and responses occur similarly across routes of exposure. This issue was examined by EPA (2010) during the development of PAH RPFs. EPA reported that animal bioassays indicate some support for the idea that absorption of PAHs occurs via oral, inhalation and dermal routes, and that the relative potencies of PAHs are on the same order of magnitude. However, EPA noted that few studies were available for comparison among pathways for any one of the PAHs studied. Although BURA contains many of the PAHs assessed by EPA, many classes of PAHs within the mixture have not been examined by either route of exposure. Therefore, existing information is limited to characterize the uncertainty for extrapolating between skin and lung exposures.

Dose–response of uncharacterized PAHs: The evaluation of whole-mixture exposures (including all chemical components) enables direct assessment of the complete mixture. This is useful for evaluating the toxicity of BURA or other asphalts due to the complex suite of PAHs known to exist within the mixture. The analyses completed herein allow for characterization of any additive, synergistic or antagonistic interactions among asphalt constituents. Generally, chemical characterization of BURA and other paving asphalts has focused on well-known carcinogenic PAHs (e.g. B[*a*]P). However, this group of PAHs may not fully explain the tumor response in animal studies (IARC 2013a; Niemeier et al. 1988; Sivak et al. 1997). Limited *in vitro* or *in vivo* data are available to describe the carcinogenic activity of other classes of PAHs (e.g. N- and S-substituted heterocyclic PAHs). IARC (2013b) summarized the evidence for a subset of heterocyclic PAHs and generally found no to limited evidence for carcinogenic activity. While the assessment of the whole mixture (and fluorescence measures) incorporates these compounds, the degree to which other classes of PAHs contribute to the tumor incidence observed in the dermal studies discussed in the "Quantitative dose–response assessment" section is uncertain. However, complete toxicological profiles for every constituent PAH are not critical, because the whole-mixture information captures their contributions to overall cancer risks.

B[a]P potency: Some (but not all) of our methods rely on assessing a relative potency for the tested asphalt fume condensate material and for B[*a*]P when tested in a similar system. This potency relative to B[*a*]P can then be used to render an exposure to BURA fumes into a B[*a*]P equivalent, which in turn enables assessment of its risk potential using an externally supplied carcinogenic potency for B[*a*]P. To make our arguments specific, we have used the latest draft values for B[*a*]P unit risks from the EPA (2014a, b), but these EPA values are drafts that may change, and their soundness has been questioned in public comments received on its draft by EPA. Our ratios of potency of BURA to B[*a*]P do not depend on the draft EPA unit risks, but translating them into BURA-specific unit risks does, and a revised view of B[*a*]P 's potency would change our projections of risks from BURA-fume exposure accordingly.

Conclusions

IARC determined that occupational exposure to oxidized asphalt and its emissions during roofing was probably carcinogenic to humans (Group 2A) based on limited evidence in humans, limited evidence in animals for oxidized asphalts, and sufficient evidence in animals for fume condensates of oxidized asphalts (IARC 2013a). This position raises the question about what levels of exposure are encountered by roofing workers that might cause cancer risks that would be of concern.

To investigate this question, we considered the available information to create a basis for estimating potential roofer cancer risks from exposure to BURA. Further, we developed a set of alternative estimates of cancer potency that we then combined with exposure information to estimate these potential risks. Concurrently, we also considered potential co-exposures to other contaminants and analyzed how these confounders affect the risk estimates. Roofers are known to have other exposures that affect their overall cancer risk, including sunlight, materials containing coal tar or asbestos in old roofing that may be removed before BURA application, and others (e.g. tobacco smoke or diesel emissions).

Compared to other PAH-containing mixtures for which cancer risks have been considered, the carcinogenic PAH content of BURA is quite low (e.g. as compared to coal tar). Moreover, the complement of particular PAHs that have been examined in different BURA sources and fume-generation methods is consistent among cases, with modest variation attributable to different source material, and some variation attributable to the temperature at which fumes are generated. This analysis justifies BURA fumes as a complex mixture that is consistent enough in composition that its potential toxicity effects can be evaluated by testing the mixture itself, rather than its components. Moreover, the differences between BURA and other PAH-containing mixtures suggest that a BURA-specific analysis of potential cancer risks is not only possible, but also warranted.

Regarding epidemiology studies of roofers, the information on skin cancer is very limited, and, as a whole, studies of lung cancer do not consistently demonstrate elevated risks. Most existing studies are hampered by lack of study-specific quantitative exposure information and by confounding by co-exposures to coal tar and/or asbestos that roofers encounter during the processes of removing old roofs. Nonetheless, we concluded that the small RRs that were estimated from some of these studies indicate that the possible magnitude of lung cancer risk for roofers with typical levels of exposure must be low and may be partially or wholly attributable to confounders (and so not necessarily attributable in whole or in part to BURA exposure).

In contrast to the epidemiology database, animal bioassay data on exposures to BURA are suitable for skin tumor evaluation (in skin-painting studies of fume condensate) but are more meager for inhalation (for which the only study available investigated fumes from paving asphalt rather than roofing asphalt). Available bioassays generally evaluate risks from exposure to B[*a*]P in parallel, either as part of the same experiment or, in the case of inhalation, as a separate study conducted according to the same methods. This fact opens the possible line of analysis to evaluate tumorigenic potency of BURA relative to that of B[*a*]P, and then to use existing route-specific estimates of human B[*a*]P cancer risks (such as those recently proposed by EPA) to project the risks of BURA exposure in terms of B[*a*]P equivalents. We used this approach in addition to using the bioassay data as direct observation and measurement of the BURA Type III mixture *per se*.

As a whole, the bioassay data show that tumors arising from BURA exposure have marked time dependence, with incidence rising as a power of time. This means that determining potency depends on the timepoint chosen for the evaluation, which could bias results when animals die or are removed from testing during the course of a chronic-exposure experiment. Fortunately, we have been able to analyze and correct for this effect using information on time of tumor appearance in individual animals and time-to-tumor modeling.

Many bioassays are of limited utility because they examine only one dose level, and often a high one that produces a substantial response. Using such information to make inferences about lower exposure levels (such as roofers would experience) requires establishing dose–response relationships. The parallel studies on B[*a*]P reveal a very nonlinear dose–response pattern. We were able to use information on B[*a*]P's dose–response shape, along with assumptions about the consistency of carcinogenic mode of action among PAHs, to make inferences about lower doses of the BURA exposures, even in single-dose studies. The time-to-tumor correction described above, and the analysis of parallel dose–response curves, enabled us to define relative potency (between BURA fractions, from one BURA sample to another, or between BURA and B[*a*]P) consistently, in a way that is not specific to dose levels or time points. Our generalizable approach makes such analysis appropriate for developing a general human cancer potency estimate that can be applied at whatever dose levels or time durations are of interest.

We quantitatively examined potential cancer risks using several different methods based on information extrapolated from the animal bioassay data. Using estimates based on measured field exposures and potency based on field-collected BURA fume, the risk estimates suggest a modest level of risk (10^{-4}). However, estimating potential dermal cancer risks based on laboratory-generated fume condensates or asphalt fume exposures, including coal tar, yields higher risk estimates. The field-based risks estimates are generally consistent with epidemiological data (i.e. no

significant increase in skin cancers). Because our estimate of B[*a*]P skin cancer risk is similar to that of EPA, our approach of estimating BURA fume skin cancer risk relative to that of B[*a*]P, and then applying the EPA skin cancer potency for B[*a*]P, does not differ substantially from the direct method (i.e. risk estimates within the same range). For inhalation risk and lung cancer, the available approaches all have limitations. The single rat study focused on paving asphalt fumes, which may differ from BURA fumes, and that study showed no evident risk.

We acknowledge that there are potential uncertainties within the underlying data and assumptions used to generate dose–response and exposure estimates. Some of these uncertainties could result in risk estimates biased either high or low. However, we also note that each of the risk estimates (based on animal bioassays and typical field exposures) for both lung and skin cancer (estimated from several possible methodological approaches) yield similar ranges of risk levels for roofing workers. Thus, the results from these approaches generally converge on a modest risk range, which tempers these uncertainties. That is, despite the challenges that make firm quantification of cancer risks to workers difficult, our analyses suggest that projections from the available methods indicate generally low (and under most regulatory schemes, acceptably low) risk levels that do not contradict the negative findings of epidemiological studies.

In conclusion, in accordance with EPA's (2000) mixtures guidance, we conducted an analysis of existing data for the primary mixture of interest (BURA). We have demonstrated that compositional variations of BURA and biological responses (for various fractions or the whole mixture) from several animal bioassays yield generally consistent results. In this regard, the toxicity and risk assessment results based on whole mixtures are preferable to component-based, e.g. risk estimates (e.g. relative potency approaches). Response metrics (mutagenicity, carcinogenicity, and potency indices) were explored and are well correlated with fluorescence measurements (i.e. measuring biologically active multi-ringed PAHs). Each of these indices showed a consistently decreasing response with decreasing fluorescence. Thus, to the extent that other asphalt mixtures are comparable to BURA, the fluorescence metric can be used qualitatively to screen mixtures for relative carcinogenic potency. This method would provide an alternative to the component-based approach, which depends on, i.e. B[*a*]P equivalents and is less reliable for asphalt mixtures. However, application of these methods would be limited to mixtures with composition and physical properties sufficiently similar to those of BURA (e.g. PAH content, distillation temperature or additives). Therefore, the results we report here would be most applicable to other roofing or paving asphalts and would be less reliable for PAH mixtures with distinct properties that differ from those of the mixtures assessed in this report. This said, we hope that the approaches we have taken to estimate potential quantitative risks from BURA fumes will be applied to evaluation of other PAH mixtures or to other assessments in which relative potency of components is at issue.

Declaration of interest

The authors are employed by Gradient Corporation, a private environmental consulting firm. The work reported in this paper was conducted during the normal course of employment. This paper was prepared with financial support from the Asphalt Roofing Environmental Council (AREC) and was reviewed by members of AREC while in preparation. None of the authors of this review participated in the IARC carcinogenic hazard review of asphalts and bitumens nor have the authors participated in any regulatory or legal proceedings related to the material covered in this review. Gradient as a company has worked for a variety of clients on some of the broader issues of assessing the human health risks of exposure to mixtures of polycyclic hydrocarbons, though not on the specific issues, methodologies or data considered in this review. The authors have the sole responsibility for the writing and contents of this paper.

References

Agency for Toxic Substances and Disease Registry (ATSDR). Toxicological profile for polycyclic aromatic hydrocarbons (PAHs) (update). Prepared by Research Triangle Institute. NTIS PB95-264370. Atlanta, GA; August 1995.

American Conference of Governmental Industrial Hygienists (ACGIH). Documentation for asphalt (bitumen) fumes. In: Documentation of the threshold limit values and biological exposure indices. 7th ed. Cincinnati, OH; 2001. 22 pp.

American Petroleum Institute (API). Asphalt category analysis and hazard characterization. Petroleum HPV testing group. Submitted to US EPA. Washington, DC; July 14, 2009.

Armstrong B, Hutchinson E, Unwin J, Fletcher T. Lung cancer risk after exposure to polycyclic aromatic hydrocarbons: a review and meta-analysis. Environ Health Perspect 2004;112:970–978.

Arthur D. Little, Inc. Roofing asphalts, pitch and UVL carcinogenesis. Report to National Institute for Occupational Safety and Health (NIOSH). Cambridge, MA; November 1981.

Arthur D. Little, Inc. Assessment of the cocarcinogenic/promoting activity of asphalt fumes. Report to National Institute for Occupational Safety and Health (NIOSH). Cambridge, MA; December 1989.

Asphalt Roofing Manufacturers Association (ARMA, Washingtom, DC), Bitumen Waterproofing Association (BWA, Nottingham UK), National Roofing Contractors Association (NRCA, Rosemont, IL), Roof Coatings Manufacturers Association (RCMA, Washington, DC). The

Bitumen Roofing Industry – a global perspective: production, use, properties, specifications and occupational exposure. 2nd ed. 2011.

Atillasoy ES, Elenitsas R, Sauter ER, Soballe PW, Herlyn M. UVB induction of epithelial tumors in human skin using a RAG-1 mouse xenograft model. J Invest Dermatol 1997; 109:704–709.

Atillasoy ES, Seykora JT, Soballe PW, Elenitsas R, Nesbit M, Elder DE, Montone KT, et al. UVB induces atypical melanocytic lesions and melanoma in human skin. Am J Pathol 1998;152:1179–1186.

Axelson O, Steenland K. Indirect methods of assessing the effects of tobacco use in occupational studies. Am J Ind Med 1988;13:105–118.

Balmain A, Harris CC. Carcinogenesis in mouse and human cells: parallels and paradoxes. Carcinogenesis 2000;21:371–377.

Battelle. Multistage Weibull time-to-tumor model in EPA's benchmark dose software (BMDS): methodology description. Report to US EPA, National Center for Environmental Assessment (NCEA). Columbus, OH; September 29, 2010.

Berking C, Takemoto R, Binder RL, Hartman SM, Ruiter DJ, Gallagher PM, Lessin SR, Herlyn M. Photocarcinogenesis in human adult skin grafts. Carcinogenesis 2002;23:181–187.

Boffetta P, Burstyn I, Partanen T, Kromhout H, Svane O, Langard S, Jarvholm B, et al. IARC epidemiological study of cancer mortality among European asphalt workers. IARC Internal Report No. 01/003. October. Lyon, France: International Agency for Research on Cancer (IARC); 2001.

Boffetta P, Burstyn I, Partanen T, Kromhout H, Svane O, Langard S, Jarvholm B, et al. Cancer mortality among European asphalt workers: an international epidemiological study. I. Results of the analysis based on job titles. Am J Ind Med 2003a;43:18–27.

Boffetta P, Burstyn I, Partanen T, Kromhout H, Svane O, Langård S, Järvholm B, et al. Cancer mortality among European asphalt workers: an international epidemiological study. II. Exposure to bitumen fume and other agents. Am J Ind Med 2003b;43:28–39.

Bolliet C, Juery C, Thiebaut B. Impact of oxidation process on polycyclic aromatic hydrocarbon (PAH) content in bitumen. J Occup Environ Hyg 2013;10:435–445.

Bosetti C, Boffetta P, La Vecchia C. Occupational exposures to polycyclic aromatic hydrocarbons, and respiratory and urinary tract cancers: a quantitative review to 2005. Ann Oncol 2007;18:431–446.

Brandt H, De Groot P. Emission and composition of fumes from current bitumen types. In: Eurasphalt & Eurobitume Congress 1996, Strasbourg, France. 1996. 10 pp.

Brandt H, De Groot P, Molyneux MKB, Tindle PE. Sampling and analysis of bitumen fumes. Ann Occup Hyg 1985;29:27–80.

Bronaugh RL, Collier SW, Macpherson SE, Kraeling ME. Influence of metabolism in skin on dosimetry after topical exposure. Environ Health Perspect 1994;102:71–74.

Brown RP, Delp MD, Lindstedt SL, Rhomberg LR, Beliles RP. Physiological parameter values for physiologically based pharmacokinetic models. Toxicol Ind Health 1997;13:407–484.

Burstyn I, Boffeta P, Kauppinen T, Heikkilä P, Svane O, Partanen T, Stücker I, et al. Estimating exposures in the asphalt industry for an international epidemiological cohort study of cancer risk. Am J Ind Med 2003;43:3–17.

Burstyn I, Kromhout H, Kauppinen T, Heikkila P, Boffetta P. Statistical modelling of the determinants of historical exposure to bitumen and polycyclic aromatic hydrocarbons among paving workers. Ann Occup Hyg 2000;44:43–56.

Cavallari JM, Osborn LV, Snawder JE, Kriech AJ, Olsen LD, Herrick RF, Mcclean MD. Predictors of airborne exposures to polycyclic aromatic compounds and total organic matter among hot-mix asphalt paving workers and influence of work conditions and practices. Ann Occup Hyg 2012a;56:138–147.

Cavallari JM, Zwack LM, Lange CR, Herrick RF, Mcclean MD. Temperature-dependent emission concentrations of polycyclic aromatic hydrocarbons in paving and built-up roofing asphalts. Ann Occup Hyg 2012b;56:148–160.

Clark CR, Burnett DM, Parker CM, Arp EW, Swanson MS, Minsavage GD, Kriech AJ, et al. Asphalt fume dermal carcinogenicity potential: I. Dermal carcinogenicity evaluation of asphalt (bitumen) fume condensates. Regul Toxicol Pharmacol 2011;61:9–16.

Dong W, Vaughan P, Sullivan K, Fletcher T. Mortality study of construction workers in the UK. Int J Epidemiol 1995;24:750–757.

dos Santos Silva I. (1999). Cancer epidemiology: principles and methods. Lyon: International Agency for Research on Cancer (IARC). Available from: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf [last accessed 1 October 2015].

Engholm G, Englund A, Linder B. Mortality and cancer incidence in Swedish road paving asphalt workers and roofers. Health Environ 1991;1:62–68.

EPA. Risk assessment guidance for superfund. Vol. I: Human health evaluation manual (Part A) (Interim Final). EPA-540/1-89-002. Office of Emergency and Remedial Response. Washington, DC; December 1989.

EPA. Draft report: a cross-species scaling factor for carcinogen risk assessment based on equivalence of $mg/kg^{3/4}$/day. Fed Reg 1992;57:24152–24173.

EPA. Provisional guidance for quantitative risk assessment of polycyclic aromatic hydrocarbons (final). EPA/600/R-93/089. Cinncinnati, OH; 1993.

EPA. Supplementary guidance for conducting health risk assessment of chemical mixtures. EPA/630/R-00/002. Risk Assessment Forum Technical Panel; Office of Research and Development – NCEA, Office of Pesticide Programs, Office of Pollution Prevention and Toxics, Office of Water; August 2000.

EPA. Risk assessment guidance for superfund (RAGS). Vol. I: Human health evaluation manual (Part E, Supplemental guidance for dermal risk assessment) (final). EPA/540/R/99/005. Office of Superfund Remediation and Technology Innovation; July 2004.

EPA. Guidelines for carcinogen risk assessment (final) EPA/630/P-03/001F. Risk Assessment Forum. Washington, DC; March 2005.

EPA. Development of a relative potency factor (RPF) approach for polycyclic aromatic hydrocarbon (PAH) mixtures – in support of summary information on the integrated risk information system (IRIS). EPA/635/R-08/012A. Washington, DC; February 2010.

EPA. Recommended use of body weight3/4 as the default method in derivation of the oral reference dose (final). EPA/100/R11/0001. Risk Assessment Forum. Washington, DC; 2011.

EPA. Benchmark dose technical guidance. EPA/100/R-12/001. Risk Assessment Forum. Washington, DC; June 2012.

EPA. Toxicological review of benzo(*a*)pyrene (CASRN 50-32-8) in support of summary information on the integrated risk information system (IRIS). EPA/635/R-14/312a. National Center for Environmental Assessment (NCEA). Washington, DC; September 2014a.

EPA. Toxicological review of benzo(*a*)pyrene (CASRN 50-32-8) in support of summary information on the integrated risk information system (IRIS). EPA/635/R-14/312b. National Center for Environmental Assessment (NCEA). Washington, DC; September 2014b.

Fayerweather WE. Meta-analysis of lung cancer in asphalt roofing and paving workers with external adjustment for confounding by coal tar. J Occup Environ Hyg 2007;4:175–200.

Feuston MH, Low LK, Hamilton CE, Mackerer CR. Correlation of systemic and developmental toxicities with chemical component classes of refinery streams. Fundam Appl Toxicol 1994;22:622–630.

Fogleman EV, Eliot M, Michaud DS, Nelson HH, Mcclean MD, Langevin SM, Kelsey KT. Occupational asphalt is not associated with head and neck cancer. Occup Med 2015;65:570–573.

Freeman JF, Schreiner CA, Beazley S, Burnett DM, Clark CR, Mahagaokar S, Parker CM, et al. Asphalt fume dermal carcinogenicity potential: II. Initiation-promotion assay of Type III built-up roofing asphalt. Regul Toxicol Pharmacol 2011;61:17–22.

Fuhst R, Creutzenberg O, Ernst H, Hansen T, Pohlmann G, Preiss A, Rittinghausen S. 24 Months inhalation carcinogenicity study of bitumen fumes in Wistar (WU) Rats. J Occup Environ Hyg 2007;4:20–43.

Gamble JF, Nicolich MJ, Barone NJ, Vincent WJ. Exposure-response of asphalt fumes with changes in pulmonary function and symptoms. Scand J Work Environ Health 1999;25:186–206.

Gennings C, Carter WH, Carchman RA, Teuschler LK, Simmons JE, Carney EW. A unifying concept for assessing toxicological interactions: changes in slope. Toxicol Sci 2005;88:287–297.

Graem N. Epidermal changes following application of 7,12-dimethylbenz(*a*)anthracene and 12-*O*-tetradecanoylphorbol-13-acetate to human skin transplanted to nude mice studied with histological species markers. Cancer Res 1986;46:278–284.

Haddad S, Charest-Tardif G, Krishnan K. Physiologically based modeling of the maximal effect of metabolic interactions on the kinetics of components of complex chemical mixtures. J Toxicol Environ Health A 2000;61:209–223.

Hammond EC, Selikoff IJ, Lawther PL, Seidman H. Inhalation of benzpyrene and cancer in man. Ann NY Acad Sci 1976;271:116–124.

Hansen ES. Cancer incidence in an occupational cohort exposed to bitumen fumes. Scand J Work Environ Health 1989;15:101–105.

Hansen ES. Mortality of mastic asphalt workers. Scand J Work Environ Health 1991;17:20–24.

Hatjian BA, Edwards JW, Harrison J, Williams FM, Blain PG. Ambient, biological, and biological effect monitoring of exposure to polycyclic aromatic hydrocarbons (PAHs). Toxicol Lett 1995;77:271–279.

Hatjian B, Edwards J, Williams F, Harrison J, Blain P. Risk assessment of occupational exposure to bitumen fumes in the road paving and roofing industries. J Occup Health Saf 1997;13:65–78.

Heinrich U, Roller M, Pott F. Estimation of a lifetime unit lung cancer risk for benzo(*a*)pyrene based on tumour rates in rats exposed to coal tar/pitch condensation aerosol. Toxicol Lett 1994;72:155–161.

Hicks JB. Asphalt industry cross-sectional exposure assessment study. Appl Occup Environ Hyg 1995;10:840–848.

Hrubec Z, Blair AE, Rogot E, Vaught J. Mortality risks by occupation among U.S. veterans of known smoking status (1954–1980). Bethesda, MD: US Department of Health and Human Services, Public Health Service, National Institutes of Health (NIH Publication No. 92-3407); 1992.

International Agency for Research on Cancer (IARC). IARC monographs on the evaluation of carcinogenic risks to humans. Vol. 103: Bitumens and bitumen emissions. IARC Monograph No. 103. Lyon, France: International Agency for Research on Cancer, World Health Organization (WHO); 2013a.

International Agency for Research on Cancer (IARC). IARC monographs on the evaluation of carcinogenic risks to humans. Vol. 103: Some N- and S-heterocyclic polycyclic aromatic hydrocarbons. IARC Monograph No. 103. Lyon/Geneva: International Agency for Research on Cancer/World Health Organization (WHO); 2013b. Available from: http://monographs.iarc.fr/ENG/Monographs/vol103/mono103-002.pdf [last accessed 1 October 2015].

Inter-Organization Programme for the Sound Management of Chemicals (IOMC). Asphalt (Bitumen). Concise International Chemical Assessment Document 59. Geneva, Switzerland: World Health Organization (WHO); 2004.

Jarvis IW, Dreij K, Mattsson A, Jernstrom B, Stenius U. Interactions between polycyclic aromatic hydrocarbons in complex mixtures and implications for cancer risk assessment. Toxicology 2014;321:27–39.

Kao J, Hall J, Helman G. *In vitro* percutaneous absorption in mouse skin: influence of skin appendages. Toxicol Appl Pharmacol 1988;94:93–103.

Kao J, Patterson FK, Hall J. Skin penetration and metabolism of topically applied chemicals in six mammalian species, including man: an *in vitro* study with benzo(a)pyrene and testosterone. Toxicol Appl Pharmacol 1985;81:502–516.

Kappes U, Schliemann-Willers S, Bankova L, Heinemann C, Fischer TW, Ziemer M, Schubert H, et al. The quality of human skin xenografts on SCID mice: a noninvasive bioengineering approach. Br J Dermatol 2004;151:971–976.

Knafla A, Petrovic S, Richardson M, Campbell J, Rowat C. Development and application of a skin cancer slope factor for exposures to benzo[a]pyrene in soil. Regul Toxicol Pharmacol 2011;59:101–110.

Kriech AJ, Emmel C, Osborn LV, Breuer D, Redman AP, Hoeber D, Bochmann F, Ruehl R. Side-by-side comparison of field monitoring methods for hot bitumen emission exposures: the German IFA Method 6305, U.S. NIOSH Method 5042, and the Total Organic Matter Method. J Occup Environ Hyg 2010;7:712–725.

Kriech A, Kurek J, Osborn L, Blackburn G, Fehsenfeld F. Bio-directed fractionation of laboratory-generated asphalt fumes: relationship between composition and carcinogenicity. Polycycl Aromat Compd 1999;14:189–199.

Kriech AJ, Osborn LV, Trumbore DC, Kurek JT, Wissel HL, Rosinski KD. Evaluation of worker exposure to asphalt roofing fumes: influence of work practices and materials. J Occup Environ Hyg 2004a;1:88–98.

Kriech AJ, Osborn LV, Wissel HL, Kurek JT, Sweeney BJ, Peregrine CJG. Total versus inhalable sampler comparison study for the determination of asphalt fume exposures within the road paving industry. J Environ Monit 2004b;6:827–833.

Kriech AJ, Osborn LV, Wissel HL, Redman AP, Smith LA, Dobbs TE. Generation of bitumen fumes using two fume generation protocols and comparison to worker industrial hygiene exposures. J Occup Environ Hyg 2007;4:6–19.

Leigh JP. Occupations, cigarette smoking, and lung cancer in the epidemiological follow-up to the NHANES I and the California Occupational Mortality Study. Bull NY Acad Med 1996;73:370–397.

Machado ML, Beatty PW, Fetzer JC, Glickman AH, Mcginnis EL. Evaluation of the relationship between PAH content and mutagenic activity of fumes from roofing and paving asphalts and coal tar pitch. Fundam Appl Toxicol 1993;21:492–499.

McClean MD, Kelsey KT, Sison JD, Quesenberry CP, Wrensch MR, Wiencke JK. A casecontrol study of asphalt and tar exposure and lung cancer in minorities. Am J Ind Med 2011;54:811–818.

Mcclean MD, Rinehart RD, Ngo L, Eisen EA, Kelsey KT, Herrick RF. Inhalation and dermal exposure among asphalt paving workers. Ann Occup Hyg 2004;48:663–671.

Mcclean MD, Rinehart RD, Sapkota A, Cavallari JM, Herrick RF. Dermal exposure and urinary 1-hydroxypyrene among asphalt roofing workers. J Occup Environ Hyg 2007;4:118–126.

Menck HR, Henderson BE. Occupational differences in rates of lung cancer. J Occup Med 1976;18:797–780.

Milham SJ. Occupational mortality in Washington state 1950–1989. Cincinnati, OH: National Institute for Occupational Safety and Health; 1997.

Moody RP, Joncas J, Richardson M, Chu I. Contaminated soils (I): *in vitro* dermal absorption of benzo[*a*]pyrene in human skin. J Toxicol Environ Health A 2007;70:1858–1865.

Morabia A, Markowitz S, Garibaldi K, Wynder EL. Lung cancer and occupation: results of a multicentre case–control study. Br J Ind Med 1992;49:721–727.

MPI Research Inc. (MPI). A two-year dermal carcinogenicity study of roofing asphalt fume condensates in mice. Report to Asphalt Institute, Lexington, KY. Study No. 1168-003. Mattawan, MI; August 16, 2010.

Muller P, Leece B, Raha D. Scientific criteria document for multimedia standards development, polycyclic aromatic hydrocarbons (PAH). Part 1: Hazard identification and dose–response assessment. February 1997.

Mundt DJ, van Wijngaarden E, Mundt KA. An assessment of the possible extent of confounding in epidemiological studies of lung cancer risk among roofers. J Occup Environ Hyg 2007;4:163–174.

Nagao T, Golor G, Hagenmaier H, Neubert D. Teratogenic potency of 2,3,4,7,8-pentachlorodibenzofuran and of three mixtures of polychlorinated dibenzo-*p*-dioxins and dibenzofurans in mice. Problems with risk assessment using TCDD toxic-equivalency factors. Arch Toxicol 1993;67:591–597.

National Institute for Occupational Safety and Health (NIOSH). Polycyclic aromatic compounds, total (PACs). NIOSH Method 5800 (Issue 1). In: NIOSH manual of analytical methods. 4th ed. Cincinnati, OH: NIOSH; January 15, 1998. pp 1–5.

National Institute for Occupational Safety and Health (NIOSH). Hazard review: health effects of occupational exposures to asphalt. DHHS (NIOSH) Publication No. 2001–110. December 2000.

National Toxicology Program (NTP). NTP research concept: polycyclic aromatic hydrocarbons (PAHs) (draft). 2012. 14 pp. Available from: http://ntp.niehs.nih.gov/ntp/about_ntp/bsc/2012/december/pahresearchconcept_508be.pdf [last accessed 1 October 2015].

Nesnow S, Triplett LL, Slaga TJ. Mouse skin carcinogenesis: application to the analysis of complex mixtures. In: Waters MD, Sandhu SS, Lewtas J, et al., eds. Short-term bioassays in the analysis of complex environmental mixtures III. Research Triangle Park, NC; Oak Ridge, TN: Springer; 1983. pp 367–390.

Niemeier RW, Thayer PS, Menzies KT, Von Thuna P, Moss CE, Burg J. A comparison of the skin carcinogenicity of condensed roofing asphalt and coal tar pitch fumes. In: Dennis AJ, ed. Polynuclear aromatic hydrocarbons: a decade of progress: proceedings of the tenth international symposium of polynuclear aromatic hydrocarbons. Columbus (OH): Battelle Press; 1988. pp 609–647.

Nisbet ICT, Lagoy PK. Toxic equivalency factors (TEFs) for polycyclic aromatic hydrocarbons (PAHs). Regul Toxicol Pharmacol 1992;16:290–300.

Olsson A, Kromhout H, Agostini M, Hansen J, Lassen CF, Johansen C, Kjaerheim K, et al. A case–control study of lung cancer nested in a cohort of European asphalt workers. Environ Health Perspect 2010;118:1418–1424.

Osborn LV, Kurek JT, Kriech AJ, Fehsenfeld FM. Luminescence spectroscopy as a screening tool for the potential carcinogenicity of asphalt fumes. J Environ Monit 2001;3:185–190.

Parrott JL, Hodson PV, Servos MR, Huestis SL, Dixon DG. Relative potency of polychlorinated dibenzo-*p*-dioxins and dibenzofurans for inducing mixed-function oxygenase activity in rainbow trout. Environ Toxicol Chem 1995;14:1041–1050.

Partanen T, Boffetta P. Cancer risk in asphalt workers and roofers: review and meta-analysis of epidemiologic studies. Am J Ind Med 1994;26:721–740.

Pittelkow MR, Perry HO, Muller SA, Maughan WZ, O'brien PC. Skin cancer in patients with psoriasis treated with coal tar. Arch Dermatol 1981;117:465–468.

Povarov AV, Retnev VM, Etlin SN. [Characteristics of cancer morbidity among workers engaged in the manufacture of hot-laid asphalt concrete]. Gig Tr Prof Zabol 1988;7:13–15.

Putzrath RM. Estimating relative potency for receptor-mediated toxicity: reevaluating the toxicity equivalence factor (TEF) model. Regul Toxicol Pharmacol 1997;25:68–78.

Reinke G, Swanson M, Paustenbach D, Beach J. Chemical and mutagenic properties of asphalt fume condensates generated under laboratory and field conditions. Mutat Res 2000;469:41–50.

Richiardi L, Boffetta P, Simonato L, Forastiere F, Zambon P, Fortes C, Gaborieau V, Merletti F. Occupational risk factors for lung cancer in men and women: a population-based case–control study in Italy. Cancer Causes Control 2004;15:285–294.

Roelofzen JH, Aben KK, Oldenhof UT, Coenraads PJ, Alkemade HA, van de Kerkhof PC, van der Valk PG, Kiemeney LA. No increased risk of cancer after coal tar treatment in patients with psoriasis or eczema. J Invest Dermatol 2010;130:953–961.

Roy TA. *In vitro* dermal absorption of AREC asphalt fume condensate TR-A using human skin. Report to Asphalt Institute. May 17. Hilton Head, SC: Port Royal Research, LLC; 2006a.

Roy TA. *In vitro* dermal absorption of AREC asphalt fume condensate TR-B using human skin. Report to Asphalt Institute. June 17. Hilton Head, SC: Port Royal Research, LLC; 2006b.

Roy TA. *In vitro* dermal absorption of AREC asphalt fume condensate TR-C using human skin. Report to Asphalt Institute. July 14. Hilton Head, SC: Port Royal Research, LLC; 2006c.

Roy TA. *In vitro* dermal absorption of AREC asphalt fume condensate TR-D using human skin. Report to Asphalt Institute. August 31. Hilton Head, SC: Port Royal Research, LLC; 2006d.

Roy TA, Kriech AJ, Mackerer CR. Percutaneous absorption of polycyclic aromatic compounds from bitumen fume condensate. J Occup Environ Hyg 2007;4:137–143.

Roy TA, Krueger AJ, Mackerer CR, Neil W, Arroyo AM, Yang JJ. SAR models for estimating the percutaneous absorption of polynuclear aromatic hydrocarbons. In: SAR and QSAR in environmental research. Vol. 9. New York: Overseas Publishers Association; 1998. pp 171–185.

Ruhl R, Musanke U, Kolmsee K, Priess R, Zoubek G, Breuer D. Vapours and aerosols of bitumen: exposure data obtained by the German Bitumen Forum. Ann Occup Hyg 2006;50:459–468.

Schoenberg JB, Stemhagen A, Mason TJ., Patterson J, Bill J, Altman RA. Occupation and lung cancer risk among New Jersey white males. J NCI 1987;79:13–21.

Silkworth JB, Lipinskas T, Stoner CR. Immunosuppressive potential of several polycyclic aromatic hydrocarbons (PAHs) found at a Superfund site: new model used to evaluate additive interactions between benzo[*a*]pyrene and TCDD. Toxicology 1995;105:375–386.

Sivak A, Niemeier R, Lynch D, Beltis K, Simon S, Salomon R, Latta R, et al. Skin carcinogenicity of condensed asphalt roofing fumes and their fractions following dermal application to mice. Cancer Lett 1997;117:113–123.

Soballe PW, Montone KT, Satyamoorthy K, Nesbit M, Herlyn M. Carcinogenesis in human skin grafted to SCID mice. Cancer Res 1996;56:757–764.

Speight JG. The chemistry and technology of petroleum. 3rd ed. New York: Marcel Dekker; 1999.

Stern FB, Ruder AM, Chen G. Proportionate mortality among unionized roofers and waterproofers. Am J Ind Med 2000;37:478–492.

Storm JE, Collier SW, Stewart RF, Bronaugh RL. Metabolism of xenobiotics during percutaneous penetration: role of absorption rate and cutaneous enzyme activity. Fundam Appl Toxicol 1990;15:132–141.

Swaen GMH, Slangen JMM. Mortality in a group of tar distillery workers and roofers. Int Arch Occup Environ Health 1997;70:133–137.

Tallarida RJ. An overview of drug combination analysis with isobolograms. J Pharmacol Exp Ther 2006;319:1–7.

Tan YM, Clewell H, Campbell J, Andersen M. Evaluating pharmacokinetic and pharmacodynamic interactions with computational models in supporting cumulative risk assessment. Int J Environ Res Public Health 2011;8:1613–1630.

Thayer PS, Harris JC, Menzies KT, Niemeier RW. Integrated chemical and biological analysis of asphalt and pitch fumes. In: Waters MD, Sandhu SS, Lewtas J, Claxton L, Chernoff N, Nesnow S, eds. Short-term bioassays in the analysis of complex environmental mixtures III. Cincinnati, OH: Springer; 1983. pp 351–366.

Thyssen J, Althoff J, Kimmerle G, Mohr U. Inhalation studies with benzo[*a*]pyrene in Syrian golden hamsters. J Natl Cancer Inst 1981;66:575–577.

Toraason M, Hayden C, Marlow D, Rinehart R, Mathias P, Werren D, Debord DG, Reid TM. DNA strand breaks, oxidative damage, and 1-OH pyrene in roofers with coal-tar pitch dust and/or asphalt fume exposure. Int Arch Occup Environ Health 2001;74:396–404.

Trumbore DC, Osborn L, Blackburn G, Niebo R, Kriech A, Maxim LD. Effect of oxidation and extent of oxidation on biologically active PACs in asphalt products. Inhal Toxicol 2011;23:745–761.

Urano K, Katakai Y, Tokuda Y, Ueyama Y, Nomura T, Yamamoto S. Failure of genotoxic carcinogens to produce tumors in human skin xenografts transplanted to SCID mice. Carcinogenesis 1995;16:2223–2226.

van Rooij JGM, Jongeneelen FJ. Review of skin permeation hazard of bitumen fumes. J Occup Environ Hyg 2007;4:237–244.

Vineis P, Thomas T, Hayes RB, Blot WJ, Mason TJ, Pickle LW, Correa P, et al. Proportion of lung cancers in males, due to occupation, in different areas of the USA. Int J Cancer 1988;42:851–856.

Wang E, Dement JM, Lipscomb H. Mortality among North Carolina construction workers, 1988–1994. Appl Occup Environ Hyg 1999;14:45–58.

Watkins DK, Chiazze L Jr, Fryar CD, Fayerweather W. A case control study of lung cancer and non-malignant respiratory disease among employees in asphalt roofing manufacturing and asphalt production. J Occup Environ Med 2002;44:551–558.

Wolff MS, Herbert R, Marcus M, Rivera M, Landrigan PL, Andrews LR. Polycyclic aromatic hydrocarbon (PAH) residues on skin in relation to air levels among roofers. Arch Environ Health 1989;44:157–163.

Wong O. An epidemiologic mortality study of a cohort of chemical workers potentially exposed to formaldehyde, with a discussion on SMR and PMR. In: Gibson JE, ed. Formaldehyde toxicity. Washington (DC): Hemisphere Publishing Corp.; 1983. pp 256–272.

Wong O, Bailey WJ, Amsel J. Cancer mortality and incidence in mastic asphalt workers (letter). Scand J Work Environ Health 1992;18:133–135.

World Health Organization (WHO)/International Programme on Chemical Safety (IPCS). (1998). Environmental health criteria 202: selected non-heterocyclic polycyclic aromatic hydrocarbons. Environmental health criteria 202. Geneva: World Health Organization (WHO).

Zahm SH, Brownson RC, Chang JC, Davis JR. Study of lung cancer histologic types, occupation, and smoking in Missouri. Am J Ind Med 1989;15:565–579.